

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: November 17, 2005, 22:03:06 ; Search time 48.7742 Seconds  
(without alignments)  
71.366 Million cell updates/sec

Title: US-10-006-177-1

Perfect score: 43

Sequence: 1 KIMDQVQQA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesep16Dec04:\*

1: Genesep1980s:\*

2: Genesep1990s:\*

3: Genesep2000s:\*

4: Genesep2001s:\*

5: Genesep2002s:\*

6: Genesep2003as:\*

7: Genesep2003bs:\*

8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	9	5	Abg32301 HLA-A2 as
2	43	100.0	902	4	Abg09335 Novel hum
3	43	100.0	912	5	Abg71106 Human ade
4	43	100.0	2742	3	Ab23012 Human APC
5	43	100.0	2842	2	Aar63508 Adenomat
6	43	100.0	2842	5	Abg90968 Human APC
7	43	100.0	2842	7	Ad56175 Rat Prote
8	43	100.0	2843	2	Aar26052 APC gene
9	43	100.0	2843	2	Aar58634 Adenomat
10	43	100.0	2843	2	Aaw11922 Adenomat
11	43	100.0	2843	2	Aaw35392 Human ade
12	43	100.0	2843	2	Aaw38370 Human ade
13	43	100.0	2843	2	Aaw76140 Human APC
14	43	100.0	2843	2	Aaw76144 Human APC
15	43	100.0	2843	3	Aab23011 Human APC
16	43	100.0	2843	5	Abg90964 Human ade
17	43	100.0	2843	5	Abg71105 Human ade
18	43	100.0	2843	7	Ad65846 Human ade
19	43	100.0	2843	7	Adj70164 Human hea
20	43	100.0	2843	8	Adj32540 Human ade
21	43	100.0	2845	8	Ado08044 Human pol
22	43	100.0	2845	8	Ado08046 Mouse pol
23	43	100.0	2860	2	Aar63507 Adenomat
24	43	100.0	2973	2	Aaw76821 Human APC
25	43	100.0	2973	3	Aay70304 Protein u

26	43	100.0	2973	4	Aay72782
27	37	86.0	348	6	Abu36272
28	35	81.4	817	8	Ab29739
29	35	81.4	821	8	Adn20138
30	34	79.1	177	4	Aam51655
31	34	79.1	488	6	Ada89646
32	34	79.1	672	2	Aaw99657
33	34	79.1	809	8	Ad27497
34	34	79.1	810	4	Aau34090
35	34	79.1	810	8	Ado59733
36	34	79.1	810	8	AdS44620
37	34	79.1	811	6	Abu17849
38	34	79.1	813	8	AdS28068
39	34	79.1	818	4	Aau37196
40	34	79.1	818	4	Aau36831
41	34	79.1	818	6	Abu16354
42	34	79.1	818	6	Abm72479
43	34	79.1	820	5	Abb48956
44	34	79.1	820	6	Abu32805
45	34	79.1	866	2	Aaw99656

## ALIGNMENTS

### RESULT 1

ABG32301

ID ABG32301 standard; peptide; 9 AA.

XX AC ABG32301;

XX AC ABG32301;

DT 05-NOV-2002 (first entry)

XX DT 05-NOV-2002 (first entry)

DE HLA-A2 associated immunogenic peptide from human APC protein.

XX DE HLA-A2 associated immunogenic peptide from human APC protein.

KW Human; immunogen; epitope; HLA-A1; human leukocyte antigen; CTL;

KW cytotoxic Y lymphocyte; cytostatic; cancer; colorectal carcinoma;

KW ovarian carcinoma; lung carcinoma; prostate carcinoma; vaccine; tumour;

KW HLA-2; passive immunotherapy; APC; adenomatous polyposis coli.

XX OS Homo sapiens.

OS Homo sapiens.

PN WO200246416-A2.

XX WO200246416-A2.

PD 13-JUN-2002.

XX 13-JUN-2002.

PF 04-DEC-2001; 2001WO-US047290.

XX 04-DEC-2001; 2001WO-US047290.

PR 04-DEC-2000; 2000US-0251022P.

XX 04-DEC-2000; 2000US-0251022P.

PR 20-DEC-2000; 2000US-0256824P.

XX 20-DEC-2000; 2000US-0256824P.

PA (ARGO-) ARGONEX INC.

XX (ARGO-) ARGONEX INC.

PI Ramakrishna V, Ross M, Philip R;

XX Ramakrishna V, Ross M, Philip R;

DR WPI; 2002-619021/66.

XX WPI; 2002-619021/66.

PT New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte,

XX New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte,

PT and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.

XX and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.

PS Claim 1; Page 50; 60pp; English.

XX Claim 1; Page 50; 60pp; English.

CC The invention relates to an immunogen comprising an isolated polypeptide

XX The invention relates to an immunogen comprising an isolated polypeptide

CC whose amino acid sequence comprises an epitopic peptide, does not include

XX whose amino acid sequence comprises an epitopic peptide, does not include

CC MAGE 4 or MFG-E8 proteins, or consists of MAGE D protein or its

XX MAGE 4 or MFG-E8 proteins, or consists of MAGE D protein or its

CC immunologically active fragment. Also included are a polynucleotide

XX immunologically active fragment. Also included are a polynucleotide

CC encoding the immunogen or its complement, a vector comprising the

XX encoding the immunogen or its complement, a vector comprising the

CC polynucleotide, a mammalian cell comprising the vector and expressing the

XX polynucleotide, a mammalian cell comprising the vector and expressing the

CC antibody specific for the immunogen. The immunogen is useful for inducing

XX antibody specific for the immunogen. The immunogen is useful for inducing

CC a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour

XX a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour

CC cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is

XX cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is

CC useful for inducing a CTL response when administered to a subject. A

CC mammalian cell that can express the immunogen, is useful for inducing a  
 CC CTL response in vitro that is specific for a tumour cell expressing HLA-1  
 CC or HLA-2. The immunogen or cell is useful for inducing CTL for treating a  
 CC subject with cancer (carcinoma, preferably colorectal carcinoma, ovarian  
 CC carcinoma, lung carcinoma and prostate carcinoma). The immunogen is also  
 CC useful for screening and diagnostic agents, for gene screening in  
 CC patients afflicted with cancer, for screening a sample for the presence  
 CC of CTUs that specifically recognise the corresponding epitopes, as a  
 CC diagnostic tool to evaluate the efficacy of the immunotherapeutic  
 CC treatments, to prepare class I MHC (major histocompatibility class)  
 CC tetramers which are utilised in conjunction with flow cytometry to  
 CC quantitate the frequency of peptide-specific CTL that are present in a  
 CC samples of lymphocytes from an individual, and for stimulating the  
 CC production of antibodies for use in passive immunotherapy, for use as  
 CC diagnostic reagents, and for use as reagents in other processes such as  
 CC affinity chromatography. The present sequence is an immunogenic epitope  
 CC of the invention derived from human APC (adenomatous polyposis coli)  
 CC protein

XX  
 XX  
 SQ Sequence 9 AA;  
 Query Match 100.0%; Score 43; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KIMDQVQQA 9  
 Db 1 KIMDQVQQA 9  
 |||||

RESULT 2  
 ABG09335  
 ID ABG09335 standard; protein; 902 AA.  
 AC ABG09335;  
 XX  
 XX  
 DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9326.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US008631.

PF 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS73522.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX Claim 20; SEQ ID NO 39694; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 902 AA;

Query Match 100.0%; Score 43; DB 4; Length 902;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
 |||||

Db 277 KIMDQVQQA 285  
 |||||

RESULT 3  
 ABG71106  
 ID ABG71106 standard; protein; 912 AA.

XX  
 AC ABG71106;

XX 08-JAN-2003 (first entry)

XX Human adenomatous polyposis coli protein Beta-catenin binding domain.

XX Human; adenomatous polyposis coli; APC; transgenic animal; colon cancer;  
 KW developmental abnormality; anticancer drug; beta-catenin;  
 KW Wnt/Wg signalling pathway; Beta-catenin binding domain.

XX Homo sapiens.

PN EP1243646-A2.

XX 25-SEP-2002.

PF 07-MAR-2002; 2002EP-00290571.

XX 19-MAR-2001; 2001US-0276483P.

XX (COUL ) COUNCIL SCI & IND RES.

XX Bhandari P, Shashidhara LS;

PI WPI; 2002-742695/81.

XX New transgenic Drosophila containing the human colon cancer gene  
 PT Adenomatous Polyposis Coli is useful as an assay model to screen for new  
 PT drugs, particularly against colon cancer.

XX Claim 2; Page 21; 52pp; English.

XX The invention describes a transgenic Drosophila whose genome comprises  
 CC the full-length human colon cancer gene Adenomatous Polyposis Coli (APC)  
 CC allowing regulated mis-expression of the APC gene resulting in  
 CC developmental abnormalities. The transgenic flies are used to screen and  
 CC validate efficacy of anticancer drugs, to identify new target proteins  
 CC interacting with beta-catenin, genes which interact with human APC to  
 CC study the biochemical function of human APC and to identify additional  
 CC components of the Drosophila Wnt/Wg signalling pathway. In particular the  
 CC flies are used to screen potential drugs against colon cancer. This is



CC the amino acid sequence of the human adenomatous polyposis coli (APC)  
 CC protein Beta-catenin binding domain DNA encoding which is incorporated in  
 CC the transgenic flies  
 XX Sequence 912 AA;  
 SQ

Query Match 100.0%; Score 43; DB 5; Length 912;  
 Best Local Similarity 100.0%; Pred. No. 6.6; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
 |||||  
 Db 787 KIMDQVQQA 795

RESULT 4  
 AAB23012  
 ID AAB23012 standard; protein; 2742 AA.

XX AC AAB23012;

XX DT 16-JAN-2001 (first entry)

XX DE Human APC protein (splice variant 2).

XX KW APC gene; Adenomatous Polyposis Coli gene; human; chromosome 5q21;  
 KW familial adenomatous polyposis; FAP locus; Gardner's syndrome; GS;  
 KW sporadic tumour; adenoma; carcinoma; cancer; lung; breast; colon; rectum;  
 KW bladder; liver; sarcoma; stomach; prostate; leukaemia; lymphoma;  
 KW tumour suppressor; anti-APC antibody; detection; diagnosis; prognosis;  
 KW genetic predisposition; drug screening; DP2.5; splice variant.

XX OS Homo sapiens.

XX PN US6114124-A.

XX PD 05-SEP-2000.

XX PF 25-MAY-1995; 95US-00450582.

XX PR 16-JAN-1991; 91GB-00000962.

XX PR 16-JAN-1991; 91GB-00000963.

XX PR 16-JAN-1991; 91GB-00000974.

XX PR 16-JAN-1991; 91GB-00000975.

XX PR 08-AUG-1991; 91US-00741940.

XX PR 12-AUG-1994; 94US-00289548.

XX PA (ICIL ) IMPERIAL CHEM IND PLC.

XX PA (UYJO ) UNIV JOHNS HOPKINS.

XX PA (UTAH ) UNIV UTAH.

XX PA (CANC-) CANCER INST.

XX PI Carlson M, Groden J, Joslyn G, Kinzler K, Markham AF, Anand R;

XX PI Albertsen H, White RL, Thliveris A, Nakamura Y, Vogelstein B;

XX PI Hedge PJ;

XX DR WPI; 2000-565003/52.

XX DR N-PSDB; AAA93450.

XX PT Detecting Adenomatous Polypoids Coli (APC) protein in a sample for  
 PT diagnosing cancers, involves contacting the sample with antibodies that  
 PT specifically bind to APC protein and detecting the complex formed.

XX PS Example 7; Fig 7A1-7W; 125pp; English.

XX CC The invention relates to a novel method for detecting Adenomatous  
 CC Polypoids Coli (APC) protein in a sample. The method involves contacting  
 CC the sample with antibodies which specifically binds to the 2843 amino  
 CC acid form of the human APC protein, or to a mutant APC protein, and  
 CC detecting an APC-antibody complex. Mutations in the APC gene play a role  
 CC in tumorigenesis, indicating that it is a tumour suppressor gene. It is  
 CC located on chromosome 5q21, which corresponds to the FAP (familial  
 CC adenomatous polyposis) locus. FAP is an autosomal dominant inherited

CC disease in which affected individuals develop hundreds to thousands of  
 CC adenomatous polyps in the colon and rectum, some of which progress to  
 CC malignancy. The FAP locus is often found to be deleted in sporadic (i.e.,  
 CC non-familial) adenomas and carcinomas, and chromosome 5q deletions have  
 CC also been observed in tumours of the lung, breast, colon, rectum,  
 CC bladder, liver, sarcomas, stomach, and prostate, and in leukaemias and  
 CC lymphomas. Although the FAP locus contains several other genes such as  
 CC FER, TBI, TB2, and MCC, it is thought that mutations in the APC gene play  
 CC a key role in the development of FAP and sporadic tumours. The method is  
 CC useful for detecting APC protein and its mutant forms in foetal tissue,  
 CC placental tissue, amniotic fluid, blood, serum or a tumour sample. The  
 CC method is useful for diagnosing or prognosing neoplastic tissue for  
 CC detecting a genetic predisposition to cancer, for detecting germline and  
 CC somatic alteration of wild-type APC genes, and for testing therapeutic  
 CC agents for the ability to suppress tumours. The present sequence  
 CC represents a 2742 amino acid splice variant of the human APC protein.  
 CC This variant is less abundant than the 2843 amino acid variant (AAB23011)

XX SQ Sequence 2742 AA;

Query Match 100.0%; Score 43; DB 3; Length 2742;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9

Db 1644 KIMDQVQQA 1652

RESULT 5

AAR63508

ID AAR63508 standard; protein; 2842 AA.

XX AC AAR63508;

XX DT 25-MAR-2003 (revised)

XX DT 23-MAY-1995 (first entry)

XX DE Adenomatous polyposis coli tumour repressor.

XX KW Adenomatous polyposis coli; tumour repressor; Gardner's syndrome;  
 KW familial adenomatous polyposis; cancer diagnosis and prognosis;  
 KW tumorigenesis suppression.

XX OS Homo sapiens.

XX PN US5352775-A.

XX PD 04-OCT-1994.

XX PF 08-AUG-1991; 91US-00741940.

XX PR 16-JAN-1991; 91GB-00000962.

XX PR 16-JAN-1991; 91GB-00000963.

XX PR 16-JAN-1991; 91GB-00000974.

XX PR 16-JAN-1991; 91GB-00000975.

XX PA (ICIL ) IMPERIAL CHEM IND.

XX PA (CANC-) CANCER INST.

XX PA (UYJO ) UNIV JOHNS HOPKINS.

XX PA (UTAH ) UNIV UTAH.

XX PI Markham AF, Hedge PJ, Anand R, Nakamura Y, Groden J, Kinzler K;  
 XX PI Thliveris A, Carlson M, Vogelstein B, Albertsen H, White RL;  
 XX PI Joslyn G;

XX DR WPI; 1994-316233/39.

XX DR N-PSDB; AAQ72297.

XX PT New human adenomatous polyposis coli DNA encoding tumour repressor - and  
 PT derived primers and probes for diagnosis, prognosis and treatment of  
 PT cancer.

XX

PS Claim 3; Col 71-84; 113pp; English.

CC AAQ72297 is a cDNA isolated from the human adenomatous polyposis coli

CC (APC) gene, it encodes the tumour suppressors described in AAR63507 and

CC AAR63508. Determination of alterations in APC or its expression products,

CC can be used for the diagnosis and prognosis of cancer e.g. colorectal,

CC lung and breast tumours; and for determining predisposition to certain

CC cancers such as familial adenomatous polyposis (FAP) and Gardner's

CC syndrome. The wild type APC gene (or a part of it) can be used

CC therapeutically to restore gene function, while primers and probes

CC derived from the cDNA (AAQ72333-400 and AAQ72541-568) can be used to

CC detect mutations. Also APC proteins or analogues can be administered to

CC animals carrying a defective gene, and epithelial cells, or transgenic

CC animals carrying a mutated APC allele are useful for detecting

CC therapeutic agents able to suppress tumorigenesis. (Updated on 25-MAR-

CC 2003 to correct PA field.)

XX

SQ Sequence 2842 AA;

Query Match 100.0%; Score 43; DB 2; Length 2842;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9

Db 1744 KIMDQVQQA 1752

RESULT 6

ABG90968

ID ABG90968 standard; peptide; 2842 AA.

XX AC ABG90968;

XX 29-NOV-2002 (first entry)

DT Human APC protein.

XX

XX Adenomatous polyposis coli; APC; human; neoplastic tissue;

KW mutation detection; tumour; cancer.

XX Homo sapiens.

OS

XX US6413727-B1.

PN

XX 02-JUL-2002.

PD

XX 25-MAY-1995; 95US-00449731.

PF

XX 16-JAN-1991; 91GB-00000962.

PR

XX 16-JAN-1991; 91GB-00000963.

PR

XX 16-JAN-1991; 91GB-00000974.

PR

XX 16-JAN-1991; 91GB-00000975.

PR

XX 08-AUG-1991; 91US-00741940.

PR

XX 12-AUG-1994; 94US-00289548.

XX

XX (UYJO ) UNIV JOHNS HOPKINS.

PA (UTAH ) UNIV UTAH.

PA (NICA-) JAPANESE FOUND CANCER RES.

PA (ZENE ) ZENECA LTD.

XX

XX Albertsen H, Anand R, Carlson M, Groden J, Hedge PJ, Joslyn G;

PI Kinzler K, Markham AF, Nakamura Y, Thliveris A, Vogelstein B;

PI White RL;

XX

XX WPI; 2002-641559/69.

DR

XX Method to aid in the diagnosis/prognosis of neoplastic tissues in humans,

PT by detecting somatic alteration of wild-type APC protein in tumor tissue

PT isolated from human, the alteration indicating neoplasia of the tissue.

XX

XX Example 15; Fig 3; 140pp; English.

XX

CC This invention relates to a novel method to aid in the diagnosis or

CC prognosis of a neoplastic tissue of a human. The method involves

CC detecting somatic alteration of wild-type adenomatous polyposis coli

CC protein in a tumour tissue isolated from a human (the alteration is

CC indicating neoplasia of the tissue). The method of the invention is

CC useful in diagnosis or prognosis of a neoplastic tissue of a human. The

CC method is useful in detection of genetic predisposition to cancer. The

CC present sequence represents a peptide sequence used in the method of the

CC invention

XX

SQ Sequence 2842 AA;

Query Match 100.0%; Score 43; DB 5; Length 2842;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9

Db 1744 KIMDQVQQA 1752

RESULT 7

ADE56175

ID ADE56175 standard; protein; 2842 AA.

XX AC ADE56175;

XX 29-JAN-2004 (first entry)

DT Rat Protein P70478, SEQ ID NO 2024.

DE

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

OS

XX WO2003016475-A2.

PN

XX 27-FEB-2003.

PD

XX 14-AUG-2002; 2002WO-US025765.

PF

XX 14-AUG-2001; 2001US-0312147P.

PR

XX 01-NOV-2001; 2001US-0346382P.

PR

XX 26-NOV-2001; 2001US-0333347P.

PR

XX (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

PA

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; P70478.

DR

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PT

XX Claim 1; Page; 1017pp; English.

PS

XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a

CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence

CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 2842 AA;

Query Match 100.0%; Score 43; DB 7; Length 2842;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
 |||||  
 Db 1743 KIMDQVQQA 1751

RESULT 8  
 AAR26052  
 ID AAR26052 standard; protein; 2843 AA.

XX AAR26052;

XX 25-MAR-2003 (revised)  
 DT 28-JAN-1993 (first entry)

XX APC gene product in familial adenomatous polyposis.

XX neoplasm; cancer; oncogene; tumour; growth; detection; diagnosis;  
 KW prognosis; treatment; sporadic colorectal carcinomas; ss.

XX Homo sapiens.

XX WO9211103-A1.

XX 06-AUG-1992.

XX 16-JAN-1992; 92WO-US000376.

XX 16-JAN-1991; 91GB-00000963.

XX 08-AUG-1991; 91US-00741940.

XX (UYJO ) UNIV JOHNS HOPKINS.

XX (ICIL ) IMPERIAL CHEM IND PLC.

XX (UTAH ) UNIV UTAH.

XX (CANC-) CANCER INST.

XX Kinzler KW, Vogelstein B, Anand R, Hedge PJ, Markham AF;

PI Albertsen H, Carlson ML, Groden JL, Joslyn G, Thliveris A, White RL;

PI Nakamura Y;

XX WPI; 1992-284685/34.

DR N-PSDB; AAQ27234.

XX Detection of somatic and germ-line alterations of human APC gene - used  
 PT to diagnose, treat and study familial adenomatous polyposis and sporadic  
 PT colorectal cancer.

XX Disclosure; Page 47; 132pp; English.

XX This sequence is encoded by the APC (Adenomatous Polyposis Coli) gene  
 CC associated with tumorigenesis, found on chromosome 5q. The sequence may  
 CC be mutated by deletions, insertions, inversions, or point mutations of the  
 CC gene. The APC gene is expressed in most normal tissues as well suggesting  
 CC that APC is a tumour suppressor. (Updated on 25-MAR-2003 to correct PN

CC field.) (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 2843 AA;

Query Match 100.0%; Score 43; DB 2; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
 |||||  
 Db 1745 KIMDQVQQA 1753

RESULT 9

AAR58634

ID AAR58634 standard; protein; 2843 AA.

XX AAR58634;

XX 25-MAR-2003 (revised)  
 DT 21-JUN-1995 (first entry)

XX Adenomatous polyposis coli protein (APC).

KW Adenomatous polyposis coli protein gene; familial adenomatous polyposis;  
 KW colorectal tumor; adenoma.

XX Homo sapiens.

XX WO9421814-A1.

XX 29-SEP-1994.

XX 21-MAR-1994; 94WO-US002987.

XX 19-MAR-1993; 93US-00034850.

XX (UYJO ) UNIV JOHNS HOPKINS.

XX Hill DE, Johnson KA, Kinzler KW, Vogelstein B;

XX WPI; 1994-317033/39.

DR N-PSDB; AAQ70633.

XX Antibodies to adenomatous polyposis coli protein - are used for detecting  
 PT mutations in the APC gene for predicting pre-disposition to cancer,  
 PT partic. colon cancer.

XX Claim 1; Page 46; 81pp; English.

XX Antibodies prepared to this protein are used for detecting mutations in  
 CC the APC gene for predicting pre-disposition to cancer, particularly colon  
 CC cancer. The Abs can detect germ line or somatic mutations indicating a  
 CC predisposition to colon cancer and possibly gastric, oesophageal,  
 CC pancreatic or small cell lung cancers. (Updated on 25-MAR-2003 to correct  
 CC PN field.)

XX Sequence 2843 AA;

Query Match 100.0%; Score 43; DB 2; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
 |||||  
 Db 1745 KIMDQVQQA 1753

RESULT 10

AAW11922

ID AAW11922 standard; protein; 2843 AA.

XX AAW11922;

XX 07-MAY-1997 (first entry)  
 DT Adenomatous polyposis coli protein.  
 DE Adenomatous polyposis coli; APC; human; antibody; mutant detection.  
 XX Homo sapiens.  
 XX JP06347459-A.  
 XX 22-DEC-1994.  
 XX 07-JUN-1993; 93JP-00136102.  
 XX 07-JUN-1993; 93JP-00136102.  
 XX (SANKO) SANKO JUNYAKU CO LTD.  
 DR WPI; 1995-070957/10.  
 XX Detection of abnormal human APC (adenomatous polyposis coli) gene - using  
 PT antibodies directed against the APC gene product.  
 XX Claim 1; Page 7-13; 13pp; Japanese.  
 XX This sequence represents the wild type human adenomatous polyposis coli  
 CC (APC) protein. This sequence, and the APC peptide fragments represented  
 CC by AAW11923-W11929 can be used in the method of the invention. The method  
 CC of the invention is for the detection of abnormal APC, using antibodies  
 CC directed against this sequence, or one of the partial peptide sequences  
 XX Sequence 2843 AA;  
 SQ

Query Match 100.0%; Score 43; DB 2; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
 Db 1745 KIMDQVQQA 1753

RESULT 11  
 AAW35392  
 ID AAW35392 standard; protein; 2843 AA.  
 XX  
 AC AAW35392;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 11-MAR-1998 (first entry)  
 XX  
 XX Human adenomatous Polyposis coli gene product.  
 XX Human; adenomatous Polyposis coli; APC; diagnosis; prognosis;  
 KW neoplastic tissue; tumour tissue; tumour repressor; mutation;  
 KW sporadic colorectal cancer; detection.  
 XX  
 XX Homo sapiens.  
 OS  
 XX US5648212-A.  
 PN  
 XX 15-JUL-1997.  
 XX  
 XX 12-AUG-1994; 94US-00289548.  
 PF  
 XX 16-JAN-1991; 91GB-00000962.  
 PR 16-JAN-1991; 91GB-00000963.  
 PR 16-JAN-1991; 91GB-00000974.  
 PR 16-JAN-1991; 91GB-00000975.  
 PR 08-AUG-1991; 91US-00741940.  
 PR 16-JAN-1991; 91GB-00000975.  
 PR 08-AUG-1991; 91US-00741940.  
 XX (NICA-) JAPANESE FOUND CANCER RES.  
 PA

PA (UTAH) UNIV UTAH.  
 PA (UYJO) UNIV JOHNS HOPKINS.  
 PA (ZENE) ZENECA LTD.  
 XX Markham A, Nakamura Y, Groden J, Carlson M, Kinzler K;  
 PI Albertsen H, Hedge PJ, Vogelstein B, Thliveris A, Anand R, White RL;  
 PI Joslyn G;  
 XX WPI; 1997-372053/34.  
 DR N-PSDB; AAT95538.  
 XX  
 XX Cancer diagnosis - by detecting mutation(s) in adenomatous polyposis coli  
 PT gene.  
 PT  
 XX Example 1; Col 33-52; 140pp; English.  
 XX  
 CC The present sequence is the human adenomatous Polyposis coli (APC) gene  
 CC product, which was used in the development of a novel method of  
 CC diagnosing or prognosing an APC gene associated neoplastic tissue. The  
 CC method comprises comparing APC gene coding sequences or mRNA in a tumour  
 CC tissue, to APC gene coding sequences or mRNA in a non-neoplastic tissue,  
 CC where a difference indicates an APC gene associated neoplasia of the  
 CC tumour tissue. APC is a tumour repressor expressed in most normal  
 CC tissues. APC mutations are found in familial adenomatous polyposis and  
 CC sporadic colorectal cancer patients. The method enables mutations to be  
 CC detected to provide an indication of predisposition to cancer. (Updated  
 CC on 25-MAR-2003 to correct PR field.)  
 XX  
 XX Sequence 2843 AA;  
 SQ

Query Match 100.0%; Score 43; DB 2; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
 Db 1745 KIMDQVQQA 1753

RESULT 12  
 AAW38370  
 ID AAW38370 standard; protein; 2843 AA.  
 XX  
 AC AAW38370;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 08-APR-1998 (first entry)  
 XX  
 XX Human adenomatous Polyposis coli gene product.  
 XX Human; adenomatous Polyposis coli; APC; diagnosis; prognosis;  
 KW neoplastic tissue; tumour tissue; tumour repressor; mutation;  
 KW sporadic colorectal cancer; detection.  
 XX  
 XX Homo sapiens.  
 OS  
 XX US5691454-A.  
 PN  
 XX 25-NOV-1997.  
 PD  
 XX 25-MAY-1995; 95US-00452654.  
 PF  
 XX 16-JAN-1991; 91GB-00000962.  
 PR 16-JAN-1991; 91GB-00000963.  
 PR 16-JAN-1991; 91GB-00000974.  
 PR 16-JAN-1991; 91GB-00000975.  
 PR 08-AUG-1991; 91US-00741940.  
 PR 12-AUG-1994; 94US-00289548.  
 XX (CANC-) CANCER INST.  
 PA (UYJO) UNIV JOHNS HOPKINS.  
 PA (ICIL) IMPERIAL CHEM IND PLC.  
 PA (UTAH) UNIV UTAH.  
 PA

XX Nakamura Y, Markham AF, Groden J, Vogelstein B, Kinzler K,  
 PI Hedge PJ, Carlson M, Anand R, Thliveris A, Albertsen H, White RL;  
 PI Joslyn G;  
 XX  
 DR WPI; 1998-017712/02.  
 DR N-PSDB; AAT96153.  
 XX  
 PT Antibodies to normal and mutant adenomatous polyposis coli proteins -  
 PT useful for detecting genetic predisposition to cancer.  
 XX  
 PS Example 1; Col 33-52; 107pp; English.  
 XX  
 CC The present sequence is the human adenomatous Polyposis coli (APC) gene  
 CC product, which was used in the development of a novel method of  
 CC diagnosing or prognosing an APC gene associated neoplastic tissue. The  
 CC method comprises comparing APC gene coding sequences or mRNA in a tumour  
 CC tissue, to APC gene coding sequences or mRNA in a non-neoplastic tissue,  
 CC where a difference indicates an APC gene associated neoplasia of the  
 CC tumour tissue. APC is a tumour repressor expressed in most normal  
 CC tissues. APC mutations are found in familial adenomatous polyposis and  
 CC sporadic colorectal cancer patients. The method enables mutations to be  
 CC detected to provide an indication of predisposition to cancer. (Updated  
 CC on 25-MAR-2003 to correct PR field.)  
 XX  
 SQ Sequence 2843 AA;  
 Query Match 100.0%; Score 43; DB 2; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KIMDQVQQA 9  
 DB 1745 KIMDQVQQA 1753  
 RESULT 13  
 AAW76140  
 ID AAW76140 standard; protein; 2843 AA.  
 AC AAW76140;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 23-NOV-1998 (first entry)  
 XX  
 DE Human APC protein #1.  
 XX  
 KW Familial adenomatous polyposis coli; APC; tumour suppressor; therapy;  
 KW chromosome 5q21; tumorigenesis; retinoblastoma; colorectal tumour; FAP;  
 KW Gardner's Syndrome; GS; predisposition.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5783666-A.  
 XX  
 PD 21-JUL-1998.  
 XX  
 PF 25-MAY-1995; 95US-00452655.  
 XX  
 PR 16-JAN-1991; 91GB-00000962.  
 PR 16-JAN-1991; 91GB-00000963.  
 PR 16-JAN-1991; 91GB-00000974.  
 PR 16-JAN-1991; 91GB-00000975.  
 PR 08-AUG-1991; 91US-00741940.  
 PR 12-AUG-1994; 94US-00289548.  
 XX  
 PA (CANC-) CANCER INST.  
 PA (UYJO) UNIV JOHNS HOPKINS.  
 PA (UTAH) UNIV UTAH.  
 PA (ZENE) ZENECA PHARM.  
 XX  
 PI Kinzler K, Joslyn G, Markham AF, Carlson M, White RL;  
 PI Thliveris A, Groden J, Anand R, Nakamura Y, Albertsen H;

PI Vogelstein B, Hedge PJ;  
 XX  
 DR WPI; 1998-427100/36.  
 DR N-PSDB; AAV56447.  
 XX  
 PT Adenomatous polyposis coli protein - useful in the treatment of cancers  
 PT associated with mutation(s) on human chromosome 5q21.  
 XX  
 PS Disclosure; Col 41-54; 102pp; English.  
 XX  
 CC This sequence represents a human familial adenomatous polyposis coli  
 CC (APC) protein from clone DP2.5. The gene for the protein is present on  
 CC human chromosome 5q21 and is also referred to as adenomatous polyposis  
 CC coli gene. It is a tumour suppressor gene, and mutations in this gene  
 CC have been associated with tumorigenesis in retinoblastoma and colorectal  
 CC tumours, and especially familial adenomatous polyposis (FAP) and  
 CC Gardner's Syndrome (GS). The protein can be used in therapy to replace  
 CC lack of native functional protein and the nucleic acids can be used for  
 CC gene therapy. The nucleic acids that encode them can also be used as  
 CC probes and primers in detection of the cancers and predisposition to it.  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 XX  
 SQ Sequence 2843 AA;  
 Query Match 100.0%; Score 43; DB 2; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KIMDQVQQA 9  
 DB 1745 KIMDQVQQA 1753  
 RESULT 14  
 AAW76144  
 ID AAW76144 standard; protein; 2843 AA.  
 AC AAW76144;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 23-NOV-1998 (first entry)  
 XX  
 DE Human APC protein #2.  
 XX  
 KW Familial adenomatous polyposis coli; APC; tumour suppressor; therapy;  
 KW chromosome 5q21; tumorigenesis; retinoblastoma; colorectal tumour; FAP;  
 KW Gardner's Syndrome; GS; predisposition.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5783666-A.  
 XX  
 PD 21-JUL-1998.  
 XX  
 PF 25-MAY-1995; 95US-00452655.  
 XX  
 PR 16-JAN-1991; 91GB-00000962.  
 PR 16-JAN-1991; 91GB-00000963.  
 PR 16-JAN-1991; 91GB-00000974.  
 PR 16-JAN-1991; 91GB-00000975.  
 PR 08-AUG-1991; 91US-00741940.  
 PR 12-AUG-1994; 94US-00289548.  
 XX  
 PA (CANC-) CANCER INST.  
 PA (UYJO) UNIV JOHNS HOPKINS.  
 PA (UTAH) UNIV UTAH.  
 PA (ZENE) ZENECA PHARM.  
 XX  
 PI Kinzler K, Joslyn G, Markham AF, Carlson M, White RL;  
 PI Thliveris A, Groden J, Anand R, Nakamura Y, Albertsen H;  
 PI Vogelstein B, Hedge PJ;  
 XX  
 DR WPI; 1998-427100/36.

```

XX Adenomatous polyposis coli protein - useful in the treatment of cancers
PT associated with mutation(s) on human chromosome 5q21.
XX
XX Disclosure; Col 63-78; 102pp; English.
XX
CC This sequence represents a human familial adenomatous polyposis coli
CC (APC) protein isolated from 87 cDNA clones. The gene for the protein is
CC present on human chromosome 5q21 and is also referred to as adenomatous
CC polyposis coli gene. It is a tumour suppressor gene, and mutations in
CC this gene have been associated with tumorigenesis in retinoblastoma and
CC colorectal tumours, and especially familial adenomatous polyposis (FAP)
CC and Gardner's Syndrome (GS). The protein can be used in therapy to
CC replace lack of native functional protein and the nucleic acids can be
CC used as probes and primers in detection of the cancers and predisposition
CC to it. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 2843 AA;
Query Match 100.0%; Score 43; DB 2; Length 2843;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KIMDQVQQA 9
Db 1745 KIMDQVQQA 1753

RESULT 15
AAB23011
ID AAB23011 standard; protein; 2843 AA.
XX
AC AAB23011;
XX
DT 16-JAN-2001 (first entry)
XX
DE Human APC protein (splice variant 1).
XX
KW APC gene; Adenomatous Polyposis Coli gene; human; chromosome 5q21;
KW familial adenomatous polyposis; FAP locus; Gardner's syndrome; GS;
KW sporadic tumour; adenoma; carcinoma; cancer; lung; breast; colon; rectum;
KW bladder; liver; sarcoma; stomach; prostate; leukaemia; lymphoma;
KW tumour suppressor; anti-APC antibody; detection; diagnosis; prognosis;
KW genetic predisposition; drug screening; DP2.5; splice variant.
XX
OS Homo sapiens.
XX
PN US6114124-A.
XX
PD 05-SEP-2000.
XX
PF 25-MAY-1995; 95US-00450582.
XX
PR 16-JAN-1991; 91GB-00000962.
PR 16-JAN-1991; 91GB-00000963.
PR 16-JAN-1991; 91GB-00000974.
PR 16-JAN-1991; 91GB-00000975.
PR 08-AUG-1991; 91US-00741940.
PR 12-AUG-1994; 94US-00289548.
XX
XX (ICIL ) IMPERIAL CHEM IND PLC.
PA (UYJO ) UNIV JOHNS HOPKINS.
PA (UTAH ) UNIV UTAH.
PA (CANC-) CANCER INST.
XX
PI Carlson M, Groden J, Joslyn G, Kinzler K, Markham AF, Anand R;
PI Albertsen H, White RL, Thliveris A, Nakamura Y, Vogelstein B;
PI Hedge PJ;
XX
XX WPI; 2000-565003/52.
DR N-PSDB; AAA93449.
XX

```

```

PT Detecting Adenomatous Polyposis Coli (APC) protein in a sample for
PT diagnosing cancers, involves contacting the sample with antibodies that
PT specifically bind to APC protein and detecting the complex formed.
XX
XX Claim 1; Fig 3A-C; 125pp; English.
XX
CC The invention relates to a novel method for detecting Adenomatous
CC Polyposis Coli (APC) protein in a sample. The method involves contacting
CC the sample with antibodies which specifically binds to the 2843 amino
CC acid form of the human APC protein, or to a mutant APC protein, and
CC detecting an APC-antibody complex. Mutations in the APC gene play a role
CC in tumorigenesis, indicating that it is a tumour suppressor gene. It is
CC located on chromosome 5q21, which corresponds to the FAP (familial
CC adenomatous polyposis) locus. FAP is an autosomal dominant inherited
CC disease in which affected individuals develop hundreds to thousands of
CC adenomatous polyps in the colon and rectum, some of which progress to
CC malignancy. The FAP locus is often found to be deleted in sporadic (i.e.,
CC non-familial) adenomas and carcinomas, and chromosome 5q deletions have
CC also been observed in tumours of the lung, breast, colon, rectum,
CC bladder, liver, sarcomas, stomach, and prostate, and in leukaemias and
CC lymphomas. Although the FAP locus contains several other genes such as
CC FFR, TBI, TB2, and MCC, it is thought that mutations in the APC gene play
CC a key role in the development of FAP and sporadic tumours. The method is
CC useful for detecting APC protein and its mutant forms in foetal tissue,
CC placental tissue, amniotic fluid, blood, serum or a tumour sample. The
CC method is useful for diagnosing or prognosing neoplastic tissue, for
CC detecting a genetic predisposition to cancer, for detecting germline and
CC somatic alteration of wild-type APC genes, and for testing therapeutic
CC agents for the ability to suppress tumours. The present sequence
CC represents a 2843 amino acid splice variant of the human APC protein.
CC This variant is more abundant than the 2742 amino acid variant (AAB23012)
XX
XX SQ Sequence 2843 AA;
Query Match 100.0%; Score 43; DB 3; Length 2843;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KIMDQVQQA 9
Db 1745 KIMDQVQQA 1753

Search completed: November 18, 2005, 01:00:49
Job time : 55.7742 secs

```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:11:06 ; Search time 8.90323 Seconds  
(without alignments)  
97.263 Million cell updates/sec

Title: US-10-006-177-1  
Perfect score: 43  
Sequence: 1 KIMDQVQA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: Pir1.\*  
2: Pir2.\*  
3: Pir3.\*  
4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	2843	1 RBHUAP	adenomatous polyo
2	43	100.0	2845	2 I49505	adenomatous polyo
3	37	86.0	348	2 S73742	cell division prot
4	35	81.4	821	1 S76330	endopeptidase Clp
5	35	81.4	839	2 A42180	endopeptidase Clp
6	35	81.4	1060	1 A40264	kinesin-related pr
7	34	79.1	128	2 T23808	hypothetical prote
8	34	79.1	449	2 F97726	signal recognition
9	34	79.1	810	1 I40508	endopeptidase Clp
10	34	79.1	813	2 G83662	class III stress r
11	34	79.1	818	2 P98919	endopeptidase [imp
12	34	79.1	820	2 A11103	endopeptidase Clp
13	34	79.1	820	2 A11465	endopeptidase Clp
14	34	79.1	1067	2 S33417	kinesin-like prote
15	33	76.7	1214	2 H75034	reverse gyrase (to
16	32	74.4	471	2 C71439	hypothetical prote
17	32	74.4	849	2 T20422	hypothetical prote
18	32	74.4	1121	2 T02764	myosin-I binding p
19	32	74.4	1624	2 C71129	probable reverse g
20	31	72.1	96	2 A75548	conserved hypoteth
21	31	72.1	103	2 E83953	hypothetical prote
22	31	72.1	108	2 C85256	Ribosomal protein
23	31	72.1	131	2 T45241	conserved hypoteth
24	31	72.1	333	1 A40548	gap junction prote
25	31	72.1	333	1 B42053	gap junction prote
26	31	72.1	382	2 P82335	NlpD-related prote
27	31	72.1	485	2 E86506	hypothetical prote
28	31	72.1	485	2 E72115	hypothetical prote
29	31	72.1	485	2 A81555	hypothetical prote

30	31	72.1	686	2 G54618	cadmium-transporti
31	31	72.1	686	2 E71895	probable heavy-met
32	31	72.1	921	2 T01775	hypothetical prote
33	31	72.1	987	2 T51360	kinesin-like heavy
34	31	72.1	1313	2 T29027	hypothetical prote
35	31	72.1	1506	2 T32909	hypothetical prote
36	31	72.1	6713	2 B89921	hypothetical prote
37	30	69.8	134	2 AF2436	hypothetical prote
38	30	69.8	208	2 S18559	2-dehydro-3-deoxy-
39	30	69.8	228	2 S76157	hypothetical prote
40	30	69.8	283	2 B72207	fth proteinase ac
41	30	69.8	589	2 I37187	arylsulfatase E (E
42	30	69.8	738	2 T00343	hypothetical prote
43	30	69.8	839	1 S71553	endopeptidase Clp
44	30	69.8	1066	1 A48669	kinesin-related pr
45	30	69.8	1181	2 C82500	lcmP-related prote

## ALIGNMENTS

### RESULT 1

RBHUAP  
adenomatous polyposis coli protein - human  
N:Alternate names: polyposis coli locus protein DP2.5; tumor suppressor APC  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C:Accession: A37261; B39658; A49319; I54271  
R:Kinzler, K.W.; Nilbert, M.C.; Su, L.K.; Vogelstein, B.; Bryan, T.M.; Levy, D.B.; Smith, S.F.; Horii, A.; Ando, H.; Miyoshi, Y.; Miki, Y.; Nishisho, I.; Nakamura, Y.  
Science 253, 661-665, 1991  
A>Title: Identification of FAP locus genes from chromosome 5q21.  
A:Reference number: A37261; MUID:91335210; PMID:1651562  
A:Accession: A37261  
A:Molecule type: mRNA  
A:Residues: 1-2843 <KIN>  
A:Cross-references: UNIPROT:P25054; GB:M74088; NID:g182396; PIDN:AAA03586.1; PID:g182397  
R:Joslyn, G.; Carlson, M.; Thliveris, A.; Albertsen, H.; Gelbert, L.; Samowitz, W.; Grod, arrington, J.; McPherson, J.; Wasmuth, J.; Le Paslier, D.; Abderrahin, H.; Cohen, D.; Le Cell 66, 601-613, 1991  
A>Title: Identification of deletion mutations and three new genes at the familial polypos A:Reference number: A39658; MUID:91330307; PMID:1678319  
A:Accession: B39658  
A:Molecule type: DNA  
A:Residues: 1-183,'L',185-969,'N',971-1308,'G',1310-1324,'SS',1326,'HSTLE',1332-1354,'P' A:Cross-references: GB:M73548; NID:g190163; PIDN:AAA60354.1; PID:g190164  
R:Miki, Y.; Nishisho, I.; Horii, A.; Miyoshi, Y.; Utsunomiya, J.; Kinzler, K.W.; Vogelst Cancer Res 52, 643-645, 1992  
A>Title: Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a A:Reference number: A44928; MUID:92119623; PMID:1310068  
A:Accession: A44928  
A:Molecule type: DNA  
A:Residues: 1506-1525 <MTK>  
A:Cross-references: GB:S78214; NID:g243541; PIDN:AAB21145.1; PID:g243542  
A>Note: sequence extracted from NCBF backbone (NCBIN:78214, NCBIF:78218)  
R:Spirio, L.; Olschwang, S.; Groden, J.; Robertson, M.; Samowitz, W.; Joslyn, G.; Gelber Cell 75, 951-957, 1993  
A>Title: Alleles of the APC gene: an attenuated form of familial polyposis.  
A:Reference number: A49319; MUID:94073973; PMID:8252630  
A:Accession: A49319  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 'G',143-171,'P',173-179 <SPI>  
A:Cross-references: GB:S67787; NID:g461061; PIDN:AAD13997.1; PID:g4261697  
R:Lambertz, S.; Ballhausen, W.G.  
Hum. Genet. 90, 650-652, 1993  
A>Title: Identification of an alternative 5' untranslated region of the adenomatous poly A:Reference number: I54271; MUID:93186137; PMID:8383094  
A:Accession: I54271  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-4 <LAM>  
A:Cross-references: GB:S56365; NID:g266243; PIDN:AAD14918.1; PID:g4262770

C;Genetics:  
A;Gene: GDB:APC  
A;Cross-references: GDB:119682; OMIM:175100  
A;Map position: 5q21.5q22  
A;Note: mutations of this gene can result in familial adenomatous polyposis or sporadic  
C;Superfamily: adenomatous polyposis coli protein  
C;Keywords: cancer; familial adenomatous polyposis; tumor suppressor  
F:1-730/Domain: leucine-rich <NTD>  
F:7-72/Region: coli #status predicted  
F:185-227/Region: coli #status predicted  
F:731-2832/Domain: serine-rich <CTD>  
F:1131-1156/Region: acidic  
F:1558-1577/Region: acidic  
F:1866-1893/Region: highly charged

Query Match 100.0%; Score 43; DB 1; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
|:|||||  
Db 1745 KIMDQVQQA 1751

RESULT 2  
149505  
adenomatous polyposis coli protein - mouse  
N;Alternate names: APC  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I49505  
R;Su. L.  
Science 256, 668-670, 1992  
A;Title: Multiple intestinal neoplasia caused by a mutation in the murine homolog of the  
A;Reference number: I49505; MUID:92263101; PMID:1350108  
A;Accession: I49505  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-2845 <RES>  
A;Cross-references: UNIPROT:Q61315; GB:M88127; NID:g191991; PIDN:AAB59632.1; PID:g191992  
C;Superfamily: adenomatous polyposis coli protein

Query Match 100.0%; Score 43; DB 2; Length 2845;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
|:|||||  
Db 1743 KIMDQVQQA 1751

RESULT 3  
S73742  
cell division protein ftsY - Mycoplasma pneumoniae (strain ATCC 29342)  
N;Alternate names: hypothetical protein A05\_orf348  
C;Species: Mycoplasma pneumoniae  
A;Variety: ATCC 29342  
C;Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C;Accession: S73742  
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirk1, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A;Reference number: S73742; MUID:97105985; PMID:8948633  
A;Accession: S73742  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-348 <HW>  
A;Cross-references: UNIPROT:P75362; EMBL:AE000040; GB:U00089; NID:g1674091; PIDN:AAB5606  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C;Genetics:  
A;Gene: ftsY  
A;Genetic code: SGC3  
C;Superfamily: docking protein

Query Match 86.0%; Score 37; DB 2; Length 348;  
Best Local Similarity 77.8%; Pred. No. 5.6;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
|:|||||  
Db 16 KLVDQVQQA 24

RESULT 4  
S76330  
endopeptidase Clp (EC 3.4.21.-) ATP-binding chain C [similarity] - Synechocystis sp. (st  
N;Alternate names: ATP-dependent Clp proteinase regulatory chain  
N;Contains: adenosinetriphosphatase [EC 3.6.1.3]  
C;Species: Synechocystis sp.  
A;Variety: PCC 6803  
C;Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: S76330  
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O., K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
S.  
A;Reference number: S74322; MUID:97061201; PMID:8905231  
A;Accession: S76330  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-821 <KAN>  
A;Cross-references: UNIPROT:Q55662; EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BA1018  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C;Function:  
A;Description: allows clpP to hydrolyze polypeptides and proteins, probably by a chaperon  
e activity; ATP hydrolysis is required for clp hydrolysis of proteins but not of smaller  
C;Superfamily: endopeptidase Clp ATP-binding chain  
C;Keywords: ATP; hydrolase; molecular chaperone; nucleotide binding; P-loop; serine prot  
F:208-215/Region: nucleotide-binding motif A (P-loop)  
F:275-280/Region: nucleotide-binding motif B  
F:545-552/Region: nucleotide-binding motif A (P-loop)  
F:613-618/Region: nucleotide-binding motif B  
F:214/Binding site: ATP (Lys) #status predicted  
F:551/Binding site: ATP (Lys) #status predicted

Query Match 81.4%; Score 35; DB 1; Length 821;  
Best Local Similarity 66.7%; Pred. No. 35;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
|:|||||  
Db 263 KIMDEIROA 271

RESULT 5  
AH2180  
endopeptidase Clp ATP-binding chain [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. Strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: AH2180  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AH2180  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-839 <KUR>  
A;Cross-references: UNIPROT:Q8YST5; GB:BA000019; PIDN:BAB74698.1; PID:g17132093; GSPDB:G  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: clpB  
C;Superfamily: endopeptidase Clp ATP-binding chain



Query Match 81.4%; Score 35; DB 2; Length 839;  
Best Local Similarity 66.7%; Pred. No. 36;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KIMDOVQQA 9  
|||||:  
Db 278 KIMDEIRQA 286  
  
RESULT 6  
A40264  
kinesin-related protein Eg5 - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 09-Jul-2004  
C:Accession: A40264; S18764  
R:Le Guellec, R.; Paris, J.; Couturier, A.; Roghi, C.; Philippe, M.  
Mol. Cell. Biol. 11, 3395-3398, 1991  
A:Title: Cloning by differential screening of a Xenopus cDNA that encodes a kinesin-related protein  
A:Reference number: A40264; MUID:91246212; PMID:1710028  
A:Accession: A40264  
A:Molecule type: mRNA  
A:Residues: 1-1060 <LEG>  
A:Cross-references: UNIPROT:P28025; EMBL:X54002; NID:G64869; PIDN:CAA37950.1; PID:G64870  
A:Note: the authors translated the codon GCT for residue 784 as Leu  
C:Superfamily: kinesin-related protein Eg5; kinesin motor domain homology  
C:Keywords: ATP; coiled coil; microtubule binding; mitosis; nucleotide binding; P-loop;  
F:12-358/Domain: kinesin motor domain homology <KMT>  
F:198-105/Region: nucleotide-binding motif A (P-loop)  
F:104/Binding site: ATP (lys) #status predicted  
  
Query Match 81.4%; Score 35; DB 1; Length 1060;  
Best Local Similarity 77.8%; Pred. No. 45;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KIMDOVQQA 9  
|||||:  
Db 1040 KIMDEVQQA 1048  
  
RESULT 7  
T23808  
hypothetical protein M28.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T23808  
R:Gardner, A.  
submitted to the EMBL Data Library, June 1995  
A:Reference number: Z19803  
A:Accession: T23808  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-128 <WIL>  
A:Cross-references: UNIPROT:O21568; EMBL:249911; PIDN:CAA90127.1; GSPDB:GN00020; CESP:M28  
A:Experimental source: clone M28  
C:Genetics:  
A:Gene: CESP:M28.5  
A:Map position: 2  
A:Introns: 52/3  
C:Superfamily: rat ribosomal protein L7a  
  
Query Match 79.1%; Score 34; DB 2; Length 128;  
Best Local Similarity 77.8%; Pred. No. 8.4;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KIMDOVQQA 9  
|||||:  
Db 21 KIMBLVQQA 29  
  
RESULT 8  
F97726  
signal recognition particle protein [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: F97726  
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; RC  
Science 293, 2093-2098, 2001  
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
A:Reference number: A97700; MUID:21442074; PMID:11557893  
A:Accession: F97726  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-449 <KUR>  
A:Cross-references: UNIPROT:Q92J55; GB:AE006914; PIDN:AAL02752.1; PID:G15619266; GSPDB:G  
C:Genetics:  
A:Gene: fff  
C:Superfamily: signal recognition particle 54X protein  
  
Query Match 79.1%; Score 34; DB 2; Length 449;  
Best Local Similarity 66.7%; Pred. No. 30;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KIMDOVQQA 9  
|||||:  
Db 359 KIMDQIDQS 367  
  
RESULT 9  
I40508  
endopeptidase Clp (EC 3.4.21.-) ATP-binding chain clpC [similarity] - Bacillus subtilis  
N:Alternate names: adenosine triphosphatase clpC; ATP-dependent Clp proteinase regulator  
N:Contains: adenosinetriphosphatase (EC 3.6.1.3)  
C:Species: Bacillus subtilis  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: I40508; S66115; I40385; H69600  
R:Msadek, T.; Kunst, F.; Rapoport, G.  
Proc. Natl. Acad. Sci. U.S.A. 91, 5788-5792, 1994  
A:Title: MecB of Bacillus subtilis, a member of the ClpC ATPase family, is a pleiotropic  
A:Reference number: I40507; MUID:94286523; PMID:8016066  
A:Accession: I40508  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-810 <RES>  
A:Cross-references: UNIPROT:P37571; EMBL:U02604; NID:G4423358; PIDN:AAA19233.1; PID:G4423  
R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.  
DNA Res. 1, 1-14, 1994  
A:Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chrom  
A:Reference number: S65967; MUID:96051385; PMID:7584024  
A:Accession: S66115  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-810 <OGA>  
A:Cross-references: EMBL:D26185; NID:G467326; PIDN:BAA05320.1; PID:G467474  
R:Kruger, E.; Volker, U.; Hecker, M.  
J. Bacteriol. 176, 3360-3367, 1994  
A:Title: Stress induction of clpC in Bacillus subtilis and its involvement in stress tol  
A:Reference number: I40385; MUID:94253006; PMID:8195092  
A:Accession: I40385  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 211-387 <RE2>  
A:Cross-references: EMBL:X75930; NID:G512561; PIDN:CAA53534.1; PID:G512562  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Ehrlich, S.D.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.  
A:Authors: P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, I

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: H69600  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-810 <KUN>  
A;Cross-references: GB:Z99104; GB:AL009126; NID:G2632267; PIDN:CAB11862.1; PID:G2632353  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: clpC; mecB  
C;Function:  
A;Description: allows clpP to hydrolyze polypeptides and proteins, probably by a chaperon e activity; ATP hydrolysis is required for Clp ATP-binding chain  
C;Superfamily: endopeptidase Clp ATP-binding chain  
C;Keywords: ATP; hydrolase; molecular chaperone; nucleotide binding; P-loop; serine prote F:208-215/Region: nucleotide-binding motif A (P-loop)  
F:275-280/Region: nucleotide-binding motif B  
F:545-552/Region: nucleotide-binding motif A (P-loop)  
F:613-618/Region: nucleotide-binding motif B  
F:613-618/Region: nucleotide-binding motif B  
F:214/Binding site: ATP (Lys) #status predicted  
F:551/Binding site: ATP (Lys) #status predicted

Query Match	79.1%;	Score 34;	DB 1;	Length 810;
Best Local Similarity	55.6%;	Pred. No. 55;		
Matches	5;	Conservative 4;	Mismatches 0;	Indels 0;
Gaps	0;			

Qy 1 KIMDOVQQA 9  
|:|:|:|  
Db 263 KVMDEIRQA 271

RESULT 10  
G83662  
class III stress response-related ATPase clpC [imported] - *Bacillus halodurans* (strain C  
C;Species: *Bacillus halodurans*  
C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C;Accession: G83662  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira Nucleic Acids Res. 28, 4317-4331, 2000  
A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and A;Reference number: A83650; MUID:20512582; PMID:11058132  
A;Accession: G83662  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-813 <STO>  
A;Cross-references: UNIPROT:Q9KG2; GB:AP001507; GB:BA000004; NID:gi0172612; PIDN:BA8038  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: clpC  
C;Superfamily: endopeptidase Clp ATP-binding chain

Query Match	79.1%;	Score 34;	DB 2;	Length 813;
Best Local Similarity	55.6%;	Pred. No. 56;		
Matches	5;	Conservative 4;	Mismatches 0;	Indels 0;
Gaps	0;			

Qy 1 KIMDOVQQA 9  
|:|:|:|  
Db 265 KVMDEIRQA 273

RESULT 11  
F89819  
endopeptidase [imported] - *Staphylococcus aureus* (strain N315)  
C;Species: *Staphylococcus aureus*  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C;Accession: F89819  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: F89819  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-818 <KUR>  
A;Cross-references: UNIPROT:Q9W78; GB:BA000018; PID:gl3700415; PIDN:BA841713.1; GSPDB:G A;Experimental source: strain N315  
C;Genetics:  
A;Gene: clpC  
C;Superfamily: endopeptidase Clp ATP-binding chain

Query Match	79.1%;	Score 34;	DB 2;	Length 818;
Best Local Similarity	55.6%;	Pred. No. 56;		
Matches	5;	Conservative 4;	Mismatches 0;	Indels 0;
Gaps	0;			

Qy 1 KIMDOVQQA 9  
|:|:|:|  
Db 263 KVMDEIRQA 271

RESULT 12  
A11103  
endopeptidase Clp ATP-binding chain C [imported] - *Listeria monocytogenes* (strain EGD-e)  
C;Species: *Listeria monocytogenes*  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: A11103  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker ok, C.; Schlueter, T.; Simoes, N.; Vazquez-Boland, J.A.; Voss, H.; Wehland, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A;Title: Comparative genomics of *Listeria* species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: A11103  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-820 <GLA>  
A;Cross-references: UNIPROT:Q8VAB6; GB:NC\_003210; PIDN:CAD00759.1; PID:gl6409597; GSPDB: A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: clpC  
C;Superfamily: endopeptidase Clp ATP-binding chain

Query Match	79.1%;	Score 34;	DB 2;	Length 820;
Best Local Similarity	55.6%;	Pred. No. 56;		
Matches	5;	Conservative 4;	Mismatches 0;	Indels 0;
Gaps	0;			

Qy 1 KIMDOVQQA 9  
|:|:|:|  
Db 262 KVMDEIRQA 270

RESULT 13  
A11465  
endopeptidase Clp ATP-binding chain C [imported] - *Listeria innocua* (strain Clip11262)  
C;Species: *Listeria innocua*  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: A11465  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A;Title: Comparative genomics of *Listeria* species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: A11465  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-820 <GLA>  
A;Cross-references: UNIPROT:Q92F43; GB:AL592022; PIDN:CAC95497.1; PID:gl6412693; GSPDB:G A;Experimental source: strain Clip11262  
C;Genetics:

A:Gene: clpC  
C:Superfamily: endopeptidase Clp ATP-binding chain

Query Match 79.1%; Score 34; DB 2; Length 820;  
Best Local Similarity 55.6%; Pred. NO. 56;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
Db 262 KVMDEIROA 270  
|:|:|:|:|:|  
|:|:|:|:|:|

RESULT 14  
S33417  
Kinesin-like protein - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: I51616; S33417  
R:Houliston, E.; Le Guellec, R.; Kress, M.; Philippe, M.; Le Guellec, K.  
Dev. Biol. 164, 147-159, 1994  
A>Title: The kinesin-related protein Eg5 associates with both interphase and spindle microtubules  
A:Reference number: I51616; PMID:94299012; PMID:8026619  
A:Accession: I51616  
A>Status: preliminary; translated from GB/EMBL/DBBJ  
A:Molecule type: mRNA  
A:Residues: 1-1067 <HOU>  
A:Cross-references: UNIPROT:Q91783; EMBL:X71864; NID:G297875; PIDN:CAA50695.1; PID:G297875  
C:Superfamily: kinesin-related protein Eg5; kinesin motor domain homology  
C:Keywords: ATP; coiled coil; microtubule binding; mitosis; nucleotide binding; P-loop  
F:19-365/Domain: kinesin motor domain homology <KWOT>  
F:105-112/Region: nucleotide-binding motif A (P-loop)  
F:111/Binding site: ATP (lys) #status predicted

Query Match 79.1%; Score 34; DB 2; Length 1067;  
Best Local Similarity 66.7%; Pred. NO. 74;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
Db 1047 KIMDEVEQS 1055  
|:|:|:|:|:|  
|:|:|:|:|:|

RESULT 15  
H75034.  
Reverse gyrase (top-rg) PAB2423 - Pyrococcus abyssi (strain Orsay)  
C:Species: Pyrococcus abyssi  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: H75034  
R:anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure  
A:Reference number: A75001  
A:Accession: H75034  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1214 <KAW>  
A:Cross-references: UNIPROT:Q9U286; GB:AJ248287; GB:AL096836; NID:G5458657; PIDN:CAB5017  
A:Experimental source: strain Orsay  
C:Genetics:  
A:Gene: top-RG; PAB2423

Query Match 76.7%; Score 33; DB 2; Length 1214;  
Best Local Similarity 55.6%; Pred. NO. 1.4e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
Db 1185 KIMDEIEEA 1193  
|:|:|:|:|:|  
|:|:|:|:|:|

Search completed: November 18, 2005, 01:09:24  
Job time : 12.0032 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:04:26 ; Search time 39.9677 Seconds  
(without alignments)  
115.311 Million cell updates/sec

Title: US-10-006-177-1

Perfect score: 43

Sequence: 1 KIMQVQQA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
43	100.0	2842	1	APC_RAT	P70478	rattus norv	
43	100.0	2843	1	APC_HUMAN	P25054	homo sapien	
43	100.0	2845	1	APC_MOUSE	Q61315	mus musculu	
4	37	86.0	348	1	FTSY_MYCPN	P75362	mycoplasma
5	35	81.4	821	2	Q55662	Q55662	synechocyst
6	35	81.4	824	2	Q8DM17	Q8dm17	synechococc
7	35	81.4	839	2	Q8YST5	Q8yst5	anabaena sp
8	35	81.4	1060	1	EG51_XENLA	P28025	xenopus lae
9	35	81.4	1067	2	Q63ZT3	Q63zt3	xenopus lae
10	34	79.1	80	2	Q7NES9	Q7nes9	gloeobacter
11	34	79.1	128	1	NHPX_CAEEL	Q21568	caenorhabdi
12	34	79.1	344	2	Q9ZEG1	Q9zeg1	bacillus an
13	34	79.1	449	1	SR54_RICCN	Q92j55	rickettsia
14	34	79.1	449	2	Q7PAN3	Q7paw3	rickettsia
15	34	79.1	810	1	CLPC_BACSU	P37571	bacillus su
16	34	79.1	810	2	Q65PD6	Q65pd6	bacillus li
17	34	79.1	811	2	Q63HB8	Q63hb8	bacillus ce
18	34	79.1	811	2	Q73FC5	Q73fc5	bacillus ce
19	34	79.1	811	2	Q81J66	Q81j66	bacillus ce
20	34	79.1	811	2	Q81VW9	Q81vw9	bacillus an
21	34	79.1	811	2	Q6HPT6	Q6hpt6	bacillus th
22	34	79.1	813	2	Q9KG22	Q9kgg2	bacillus ha
23	34	79.1	818	2	Q8NXY8	Q8nxy8	staphylococ
24	34	79.1	818	2	Q99W78	Q99w78	staphylococ
25	34	79.1	818	2	Q7A797	Q7a797	staphylococ
26	34	79.1	818	2	Q6GBW3	Q6gbw3	staphylococ
27	34	79.1	818	2	Q6GJ84	Q6gje4	staphylococ
28	34	79.1	820	2	Q8YAB6	Q8yab6	listeria mo
29	34	79.1	820	2	Q92F43	Q92f43	listeria in
30	34	79.1	820	2	Q72410	Q72410	listeria mo
31	34	79.1	825	2	Q48760	Q48760	listeria mo

32	34	79.1	1067	1	EG52_XENLA	Q91783 xenopus lae
33	34	79.1	1067	2	Q8AVR8	Q8avk8 xenopus lae
34	33	76.7	393	2	Q7VNI4	Q7vni4 haemophilus
35	33	76.7	781	2	Q7S7G4	Q7s7g4 neurospora
36	33	76.7	867	2	Q6MW22	Q6mw22 neurospora
37	33	76.7	1214	1	RCYR_PYRAB	Q9uz86 pyrococcus
38	32	74.4	135	2	Q7MPW6	Q7mpw6 vibrio vuln
39	32	74.4	135	2	Q87TA6	Q87ta6 vibrio para
40	32	74.4	135	2	Q8DDW0	Q8ddw0 vibrio vuln
41	32	74.4	365	2	Q6MZQ1	Q6mzq1 homo sapien
42	32	74.4	392	1	L7L2_HUMAN	Q9y383 homo sapien
43	32	74.4	392	1	L7L2_MOUSE	Q7fnc4 mus musculu
44	32	74.4	471	2	O23552	O23552 arabidopsis
45	32	74.4	508	2	Q6CL36	Q6cl36 kluyveromyc

## ALIGNMENTS

### RESULT 1

APC\_RAT ID - APC RAT STANDARD; PRT; 2842 AA.  
AC P70478;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Adenomatous polyposis coli protein (APC protein).  
GN Name=ApC;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Fischer 344/N; TISSUE=Brain;  
RX MEDLINE=96116966; PubMed=8561176;  
RA Toyota M., Ushijima T., Kakiuchi H., Watanabe M., Imai K., Yachi A.,  
RA Sugimura T., Nagao M.;  
RT "CDNA cloning of the rat APC gene and assignment to chromosome 18.";  
RL Mamm. Genome 6:746-748(1995).  
RN [2]  
RP MUTAGENESIS.  
RC STRAIN=Fischer 344/N, and Sprague-Dawley;  
RX MEDLINE=95148647; PubMed=7846077;  
RA Kakiuchi H., Watanabe M., Ushijima T., Toyota M., Imai K.,  
RA Weisburger J.H., Sugimura T., Nagao M.;  
RT "Specific 5'-GGGA-3'-->5'-GGA-3' mutation of the APC gene in rat colon  
tumors induced by 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:910-914(1995).  
CC -!- FUNCTION: Tumor suppressor. Promotes rapid degradation of CTNNB1  
and participates in Wnt signaling. APC activity is correlated with  
its phosphorylation state (By similarity).  
CC -!- SUBUNIT: Forms homooligomers. Associates with catenins. Binds  
axin. Interacts with the N-terminus of ARHGAP4. Found in a complex  
consisting of ARHGAP4, APC and CTNNB1 (By similarity).  
CC -!- PTM: Phosphorylated by GSK3B (By similarity).  
CC -!- SIMILARITY: Contains 7 ARM repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.ebi.ac.uk/ebis-sib.ch/announcement/>)  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).  
CC -----  
DR EMBL; D38629; BAA07609.1; --  
DR HSP; P25054; IM51.  
DR RGD; 2123; APC.  
DR InterPro; IPR009240; APC\_15aa.  
DR InterPro; IPR009234; APC\_Basic.  
DR InterPro; IPR009223; APC\_err.  
DR InterPro; IPR00938; ARM.

DR InterPro; IPR000225; Armadillo.  
DR InterPro; IPR009232; EBI binding.  
DR InterPro; IPR009224; SAMF.  
DR Pfam; PF05972; APC\_15aa; 4.  
DR Pfam; PF05956; APC\_basic; 1.  
DR Pfam; PF05923; APC\_crr; 7.  
DR Pfam; PF00514; Arm; 4.  
DR Pfam; PF05937; EBI\_binding; 1.  
DR Pfam; PF05924; SAMF; 3.  
DR SMART; SM00185; ARM; 6.  
DR PROSITE; PS0176; ARM\_REPEAT; 1.  
DR Anti-oncogene; Coiled\_coil; Phosphorylation; Repeat;  
KW Wnt signaling pathway.  
FT DOMAIN 1 728 Leu-rich.  
FT DOMAIN 1 62 Coiled coil (Potential).  
FT DOMAIN 125 260 Coiled coil (Potential).  
FT REPEAT 451 493 ARM 1.  
FT REPEAT 503 545 ARM 2.  
FT REPEAT 546 589 ARM 3.  
FT REPEAT 590 636 ARM 4.  
FT REPEAT 637 681 ARM 5.  
FT REPEAT 682 723 ARM 6.  
FT REPEAT 724 765 ARM 7.  
FT DOMAIN 739 2831 Ser-rich.  
FT DOMAIN 1130 1155 Asp/Glu-rich (acidic).  
FT DOMAIN 1556 1575 Asp/Glu-rich (acidic).  
FT DOMAIN 1864 1891 Highly charged.  
FT MUTAGEN 523 523 C->R; In an IQ-induced colon tumor.  
SQ SEQUENCE 2842 AA; 310530 MW; 3CB52EA8A34E8F47 CRC64;  
Query Match 100.08; Score 43; DB 1; Length 2842;  
Best Local Similarity 100.08; Pred. No. 20;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Ov 1 KIMDQVQQA 9  
Db 1743 KIMDQVQQA 1751  
RESULT 2  
APC\_HUMAN  
ID APC\_HUMAN STANDARD; PRT; 2843 AA.  
AC P25054; Q15162; Q15163; Q93042;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Adenomatous polyposis coli protein (APC protein).  
GN Name=APC; Synonyms=DP2.5;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91335210; PubMed=16511562;  
RA Kinzler K.W., Nibert M.C., Su L.-K., Vogelstein B., Bryan T.M.,  
RA Levy D.B., Smith K.J., Preisinger A.C., Hedge P., McKechnie D.,  
RA Finnear R., Markham A., Groffen J., Boguski M.S., Altschul S.F.,  
RA Horii A., Ando H., Miyoshi Y., Miki Y., Nishisho I., Nakamura Y.;  
RT "Identification of FAP locus genes from chromosome 5q21.";  
RL Science 253:661-665 (1991).  
RN [2]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RC TISSUE=Fetal brain;  
RX MEDLINE=91330307; PubMed=1678319; DOI=10.1016/0092-8674(81)90022-2;  
RA Joslyn G., Carlson M., Thliveris A., Albertsen H., Gelbert L.,  
RA Samowitz W., Groden J., Stevens J., Spirio L., Robertson M.,  
RA Sargeant L., Krapcho K., Wolff E., Burt R., Hughes J.P.,  
RA Warrington J., McPherson J.D., Wasmuth J., Le Fasiier D.,  
RA Aberrahim H., Cohen D., Leppert M., White R.;  
RT "Identification of deletion mutations and three new genes at the  
RT familial polyposis locus.";  
RL Cell 66:601-613 (1991).

RN [3]  
RP SEQUENCE OF 1506-1525 FROM N.A.  
RX PubMed=1310068;  
RA Miki Y., Nishisho I., Horii A., Miyoshi Y., Utsunomiya J.,  
RA Kinzler K.W., Vogelstein B., Nakamura Y.;  
RT "Disruption of the APC gene by a retrotransposal insertion of L1  
RT sequence in a colon cancer.";  
RL Cancer Res. 52:643-645 (1992).  
RN [4]  
RP ASSOCIATION WITH CATENINS.  
RX MEDLINE=94082295; PubMed=8259519;  
RA Su L.-K., Vogelstein B., Kinzler K.W.;  
RT "Association of the APC tumor suppressor protein with catenins.";  
RL Science 262:1734-1737 (1993).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 3), FUNCTION, SUBCELLULAR LOCATION,  
RP INTERACTIONS WITH ARHGEF4, AND IDENTIFICATION IN A COMPLEX WITH  
RP ARHGEF4 AND CTNNB1.  
RX PubMed=10947987;  
RA Kawasaki Y., Senda T., Ishidate T., Koyama R., Morishita T.,  
RA Iwayama Y., Higuchi O., Akiyama T.;  
RT "Asef, a link between the tumor suppressor APC and G-protein  
RT signaling.";  
RL Science 289:1194-1197 (2000).  
RN [6]  
RP DISEASE.  
RX MEDLINE=95174843; PubMed=7661930; DOI=10.1056/NEJM1995033033213302;  
RA Hamilton S.R., Liu B., Parsons R.E., Papadopoulos N., Jen J.,  
RA Powell S.M., Krush A.J., Berk T., Cohen Z., Tetu B., Burger P.C.,  
RA Wood P.A., Taqi F., Booker S.V., Petersen G.M., Offerhaus G.J.A.,  
RA Tersmette A.C., Giardiello F.M., Vogelstein B., Kinzler K.W.;  
RT "The molecular basis of Turcot's syndrome.";  
RL N. Engl. J. Med. 332:839-847 (1995).  
RN [7]  
RP DISEASE.  
RX MEDLINE=97094176; PubMed=8940264;  
RA Eccles D.M., van der Luijt R.B., Breukel C., Bullman H., Bunyan D.,  
RA Fisher A., Barber J., du Boulay C., Primrose J., Burn J., Podde R.;  
RT "Hereditary desmoid disease due to a frameshift mutation at codon 1924  
RT of the APC gene.";  
RL Am. J. Hum. Genet. 59:1193-1201 (1996).  
RN [8]  
RP DISEASE.  
RX MEDLINE=20243021; PubMed=10782927;  
RA Couture J., Mitri A., Lagace R., Smits R., Berk T., Bouchard H.-L.,  
RA Fodde R., Alman B., Bapat B.;  
RT "A germline mutation at the extreme 3' end of the APC gene results in  
RT a severe desmoid phenotype and is associated with overexpression of  
RL beta-catenin in the desmoid tumor.";  
RL Clin. Genet. 57:205-212 (2000).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 2-55.  
RX MEDLINE=20384842; PubMed=10926498; DOI=10.1006/jmbi.2000.3895;  
RA Day C.L., Alber T.;  
RT "Crystal structure of the amino-terminal coiled-coil domain of the APC  
RT tumor suppressor.";  
RL J. Mol. Biol. 301:147-156 (2000).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 1021-1035 IN COMPLEX WITH  
RP CTNNB1.  
RX MEDLINE=21564054; PubMed=11707392; DOI=10.1093/emboj/20.22.6203;  
RA Eklof Spink K., Fridman S.G., Weis W.I.;  
RT "Molecular mechanisms of beta-catenin recognition by adenomatous  
RT polyposis coli revealed by the structure of an APC-beta-catenin  
RT complex.";  
RL EMBO J. 20:6203-6212 (2001).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (1.90 ANGSTROMS) OF 2034-2049 IN COMPLEX WITH  
RP AXIN.  
RX MEDLINE=20271867; PubMed=10811618; DOI=10.1093/emboj/19.10.2270;  
RA Spink K.E., Polakis P., Weis W.I.;  
RT "Structural basis of the axin-adenomatous polyposis coli  
RT interaction.";

RL EMBO J. 19:2270-2279(2000).  
 RN [12]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=94154728; PubMed=8111410;  
 RA Nagase H., Nakamura Y.;  
 RT "Mutations of the APC (adenomatous polyposis coli) gene.";  
 RL Hum. Mutat. 2:425-434(1993).  
 RN [13]  
 RP VARIANTS FAP.  
 RX MEDLINE=9135211; PubMed=1651563;  
 RA Nishihito I., Nakamura Y., Miyoshi Y.,  
 RA Koyama K., Utsunomiya J., Baba S., Hedge P.,  
 RA Petersen G.M., Hamilton S.R., Nilbert M.C.,  
 RA Preisinger A.C., Smith K.J., Su L.-K., Kinzler K.W., Vogelstein B.;  
 RT "Mutations of chromosome 5q21 genes in FAP and colorectal  
 RT patients.";  
 RL Science 253:665-669(1991).  
 RN [14]  
 RP VARIANTS FAP.  
 RX MEDLINE=91265030; PubMed=1338904;  
 RA Miyoshi Y., Nagase H., Ando H., Ichii S.,  
 RA Miki Y., Mori T., Nakamura Y.;  
 RT "Somatic mutations of the APC gene in colorectal tumors: mutation  
 RT cluster region in the APC gene.";  
 RL Hum. Mol. Genet. 1:229-233(1992).  
 RN [15]  
 RP VARIANTS FAP.  
 RX MEDLINE=93244793; PubMed=1338691;  
 RA Nakatsuru S., Yanagisawa A., Ichii S.,  
 RA Nakamura Y., Horii A.;  
 RT "Somatic mutation of the APC gene in gastric cancer: frequent  
 RT mutations in very well differentiated adenocarcinoma and signet-ring  
 RT cell carcinoma.";  
 RL Hum. Mol. Genet. 1:559-563(1992).  
 RN [16]  
 RP VARIANT FAP TRP-1348, AND VARIANTS ASP-1118; MET-1292; VAL-1304 AND  
 RP SER-2502.  
 RX MEDLINE=93250848; PubMed=1338764;  
 RA Nagase H., Miyoshi Y., Horii A., Aoki T.,  
 RA Vogelstein B., Maher E., Ogawa M., Maruyama M., Utsunomiya J.,  
 RA Baba S., Nakamura Y.;  
 RT "Screening for germ-line mutations in familial adenomatous polyposis  
 RT patients: 61 new patients and a summary of 150 unrelated patients.";  
 RL Hum. Mutat. 1:467-473(1992).  
 RN [17]  
 RP VARIANT FAP TRP-99.  
 RX TISSUE=Peripheral blood lymphocytes;  
 RC MEDLINE=95134544; PubMed=7633149; DOI=10.1016/0959-8049(94)00294-F;  
 RA Dobbie Z., Szycher M., Huerliman R., Ammann T., Roth J.,  
 RA Mueller A., Mueller H., Scott R.J.;  
 RT "Mutational analysis of the first 14 exons of the adenomatous  
 RT polyposis coli (APC) gene.";  
 RL Eur. J. Cancer 30A:1709-1713(1994).  
 RN [18]  
 RP VARIANT FAP GLY-722.  
 RX MEDLINE=9513430; PubMed=7633331;  
 RA Stella A., Montero M., Resta N., Marchese C.,  
 RA Romio L., Pilia S., Prete F., Marenzi C., Guanti G.;  
 RT "Four novel mutations of the APC (adenomatous polyposis coli) gene in  
 RT FAP patients.";  
 RL Hum. Mol. Genet. 3:1687-1698(1994).  
 RN [19]  
 RP ERRTUM.  
 RA Stella A., Montero M., Resta N., Marchese C.,  
 RA Romio L., Pilia S., Prete F., Marenzi C., Guanti G.;  
 RL Hum. Mol. Genet. 3:1918-1918(1994).  
 RN [20]  
 RP VARIANT FAP ILE-171.  
 RX MEDLINE=97144176; PubMed=8990002;  
 RC DOI=10.1002/(SICI)1098-1004(1997)9:1<7::AID-HUMU23.3.CO;2-7;  
 RA van der Luijt R.B., Meera Khan P., Vassen H.F.A.,  
 RA van Leeuwen-Cornelisse I.S.J., Wijnen J.T., van der Klift H.M.,  
 RA Plug R.J., Griffioen G., Fodde R.;

RT "Molecular analysis of the APC gene in 105 Dutch kindreds with  
 RT familial adenomatous polyposis: 67 germline mutations identified by  
 RT DGGE, PTT, and southern analysis.";  
 RL Hum. Mutat. 9:7-16(1997).  
 RN [21]  
 RP VARIANTS COLORECTAL CARCINOMA THR-880; ILE-890 AND VAL-1508.  
 RX MEDLINE=98080146; PubMed=9419979; DOI=10.1038/sj.onc.1201668;  
 RA Miyaki M., Nishio J., Konishi M., Kikuchi-Yanoshita R., Tanaka K.,  
 RA Muraoka M., Nagato M., Chong J.-M., Koike M., Terada T., Kawahara Y.,  
 RA Fukutome A., Tomiyama J., Chuganji Y., Momoi M., Utsunomiya J.;  
 RT "Drastic genetic instability of tumors and normal tissues in Turcot  
 RT syndrome.";  
 RL Oncogene 15:2877-2881(1997).  
 RN [22]  
 RP VARIANT LYS-1307.  
 RX MEDLINE=98400248; PubMed=9731522; DOI=10.1038/1666;  
 RA Redston M., Nathanson K.L., Yuan Z.Q., Neuhausen S.L., Satagopan J.,  
 RA Wong N., Yang D., Nafa D., Abrahamson J., Ozelik H.,  
 RA Antin-Ozerkis D., Andrulis I., Daly M., Pinsky L., Schrag D.,  
 RA Gallinger S., Backack M., King M.-C., Woodage T., Brody L.C.,  
 RA Godwin A., Warner E., Weber B., Foulkes W., Offit K.;  
 RT "The APC I1307K allele and breast cancer risk.";  
 Query Match 100.0%; Score 43; DB 1; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KIMDQVQQA 9  
 DB 1745 KIMDQVQQA 1753  
 RESULT 3  
 APC MOUSE  
 ID APC MOUSE STANDARD; PRT; 2845 AA.  
 AC Q6L315; Q62044;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Adenomatous polyposis coli protein (APC protein) (mAPC).  
 GN Name=ApC;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS.  
 RC STRAIN=C57BL/6J, and CAST/EI, TISSUE=Brain;  
 RX MEDLINE=92263101; PubMed=1350108;  
 RA Su L.-K., Kinzler K.W., Vogelstein B., Preisinger A.C., Moser A.R.,  
 RA Luongo C., Gould K.A., Dove W.F.;  
 RT "Multiple intestinal neoplasia caused by a mutation in the murine  
 RT homolog of the APC gene.";  
 RL Science 256:668-670(1992).  
 RN [2]  
 RP ERRTUM.  
 RA Su L.-K., Kinzler K.W., Vogelstein B., Preisinger A.C., Moser A.R.,  
 RA Luongo C., Gould K.A., Dove W.F.;  
 RL Science 256:1114-1114(1992).  
 RN [3]  
 RP SEQUENCE OF 1-45 FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RA Dicker F., Lambertz S., Reitmair A., Ballhausen W.G.;  
 RT "The murine APC gene: alternative splicing of 5' untranslated region  
 RT segments.";  
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP ALTERNATIVE SPLICING.  
 RX MEDLINE=94061824; PubMed=8242607;  
 RA Oshima M., Sugiyama H., Kitagawa K., Taketo M.;  
 RT "APC gene messenger RNA: novel isoforms that lack exon 7.";  
 RL Cancer Res. 53:5585-5591(1993).  
 CC -1- FUNCTION: Tumor suppressor. Promotes rapid degradation of CTNNB1

CC and participates in Wnt signaling. APC activity is correlated with  
 CC its phosphorylation state (By similarity).  
 CC -!- SUBUNIT: Forms homooligomers. Associates with catenins. Binds  
 CC axin. Interacts with the N-terminus of ARHGEF4. Found in a complex  
 CC consisting of ARHGEF4, APC and CTNNB1 (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=4;  
 CC Name=1;  
 CC IsoId=O61315-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=O61315-2; Sequence=VSP\_004116;  
 CC Name=3;  
 CC IsoId=O61315-3; Sequence=VSP\_004117;  
 CC Name=4;  
 CC IsoId=O61315-4; Sequence=VSP\_004116, VSP\_004117;  
 CC -!- TISSUE SPECIFICITY: Expressed in liver, spleen, kidney, heart,  
 CC lung, brain, stomach, intestine, testis and ovary.  
 CC -!- PTM: Phosphorylated by GSK3B (By similarity).  
 CC -!- SIMILARITY: Contains 7 ARM repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M88127; AAB59632.1; -;  
 CC EMBL; U02937; AAA03443.1; -;  
 CC PIR; I49505; I49505.  
 CC HSP; P25054; IDEB.  
 CC MGD; MG1:88039; Apc.  
 CC GO; GO:0005737; C:cytoplasm; IDA.  
 CC GO; GO:0005634; C:nucleus; IDA.  
 CC GO; GO:0080113; F:beta-catenin binding; IDA.  
 CC GO; GO:0009952; P:anterior/posterior pattern formation; IMP.  
 CC GO; GO:0009798; P:axis specification; IMP.  
 CC GO; GO:0009953; P:dorsal/ventral pattern formation; IMP.  
 CC GO; GO:0016055; P:Wnt receptor signaling pathway; IDA.  
 CC InterPro; IPR009240; APC 15aa.  
 CC InterPro; IPR009234; APC basic.  
 CC InterPro; IPR009223; APC\_crr.  
 CC InterPro; IPR008938; ARM.  
 CC InterPro; IPR00225; Armadillo.  
 CC InterPro; IPR009232; EBI binding.  
 CC InterPro; IPR009224; SAMP.  
 CC Pfam; PF05972; APC 15aa; 4.  
 CC Pfam; PF05956; APC basic; 1.  
 CC Pfam; PF05923; APC\_crr; 7.  
 CC Pfam; PF0514; Arm; 4.  
 CC Pfam; PF05937; EBI binding; 1.  
 CC Pfam; PF05924; SAMP; 3.  
 CC SMART; SM00185; ARM; 6.  
 CC PROSITE; PS00176; ARM REPEAT; 1.  
 CC Alternative splicing; Anti-oncogene; Coiled coil; Phosphorylation;  
 KW Repeat; Wnt signaling pathway.  
 FT DOMAIN 1 61 Coiled coil (Potential).  
 FT DOMAIN 125 245 Coiled coil (Potential).  
 FT DOMAIN 451 728 Leu-rich.  
 FT REPEAT 451 493 ARM 1.  
 FT REPEAT 503 545 ARM 2.  
 FT REPEAT 546 589 ARM 3.  
 FT REPEAT 590 636 ARM 4.  
 FT REPEAT 637 681 ARM 5.  
 FT REPEAT 682 723 ARM 6.  
 FT REPEAT 724 765 ARM 7.  
 FT DOMAIN 739 2834 Ser-rich.  
 FT DOMAIN 1130 1156 Asp/Glu-rich (acidic).  
 FT DOMAIN 1556 1575 Asp/Glu-rich (acidic).  
 FT DOMAIN 1864 1891 Highly charged.  
 FT VARSPPLIC 243 276 Missing (in isoform 2 and isoform 4).  
 FT /FTId=VSP\_004116.

FT VARSPPLIC 310 410 Missing (in isoform 3 and isoform 4).  
 FT /FTId=VSP\_004117.  
 FT VARIANT 120 120 T -> A (in strain CAST/Ei).  
 FT VARIANT 493 493 V -> I (in strain CAST/Ei).  
 FT VARIANT 797 797 Y -> F (in strain CAST/Ei).  
 FT VARIANT 1330 1330 A -> T (in strain CAST/Ei).  
 FT VARIANT 1618 1618 A -> S (in strain CAST/Ei).  
 FT VARIANT 2294 2294 G -> A (in strain CAST/Ei).  
 FT VARIANT 2496 2496 H -> Q (in strain CAST/Ei).  
 FT VARIANT 2523 2523 T -> A (in strain CAST/Ei).  
 FT VARIANT 2813 2813 T -> S (in strain CAST/Ei).  
 SQ SEQUENCE 2845 AA; 311086 MW; 145CA73CF570A99 CRC64;  
 Query Match 100.0%; Score 43; DB 1; Length 2845;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KIMDQVQQA 9  
 Db 1743 KIMDQVQQA 1751  
 RESULT 4  
 ID\_FTSY\_MYCPN STANDARD; PRT; 348 AA.  
 AC P75362;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Cell division protein ftsy homolog.  
 GN Name=ftsY; OrderedLocusNames=MPN425; ORFNames=MP416;  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633; DOI=10.1093/nar/24.22.4420;  
 RA Himmelreich R., Hilbert H., Plagens H., Pickl E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -!- FUNCTION: Functional homolog of SRP receptor. Probably involved in  
 CC the reception and insertion of a subset of proteins at the  
 CC membrane (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).  
 CC -!- SIMILARITY: Belongs to the GTP-binding SRP family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE000040; AAB96064.1; -;  
 DR PIR; S73742; S73742.  
 DR HSP; P10121; 1FTS.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR004390; ABC\_transpt\_FtsY.  
 DR InterPro; IPR000897; SRP54.  
 DR Pfam; PF00448; SRP54; 1.  
 DR Pfam; PF02881; SRP54 N; 1.  
 DR ProDom; PD000819; SRP54; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR TIGRFAMs; TIGR00064; ftsY; 1.  
 DR PROSITE; PS00300; SRP54; 1.  
 KW Cell division; Complete proteome; GTP-binding; Membrane; RNA-binding;  
 FT Signal recognition particle. GTP (By similarity).  
 FT NP\_BIND 143 150  
 FT NP\_BIND 225 229 GTP (By similarity).  
 FT



```

FT NP BIND      289      292      GTP (8y similarity).
SQ SEQUENCE    348 AA; 38775 MW; 8A14C4DB9E1BAE29 CRC64;

Query Match      86.0%; Score 37; DB 1; Length 348;
Best Local Similarity 77.8%; Pred. No. 42;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KIMDOVQQA 9
       |:|||||
Db      16 KLVDQVQQA 24

RESULT 5
Q55662
ID Q55662      PRELIMINARY;      PRT;      821 AA.
AC Q55662;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ATP-dependent Clp protease regulatory subunit.
GN Name=clpC; OrderedLocuNames=g110020;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,
RA Horouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimp S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D64000; BAA10182.1; -.
DR F1R; S76330; S76330.
DR HSSP; P03815; 1JBK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR GO; GO:0019538; P:protein metabolism; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR001270; Chaprinin_clpA/B.
DR InterPro; IPR004176; Clp_N.
DR EMBL; D64000; BAA10182.1; -.
DR F1R; S76330; S76330.
DR HSSP; P03815; 1JBK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR GO; GO:0019538; P:protein metabolism; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR001270; Chaprinin_clpA/B.
DR InterPro; IPR004176; Clp_N.
DR InterPro; IPR001943; UvrB/C.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF02861; Clp_N; 2.
DR Pfam; PF02151; UVR; 1.
DR PRINTS; SM00382; AAA; 2.
DR PROSITE; PS00870; CLPAB_1; 1.
DR PROSITE; PS00871; CLPAB_2; 1.
DR PROSITE; PS0151; UVR; 1.
DR ATP-binding; Complete proteome; Protease.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR001270; Chaprinin_clpA/B.
DR InterPro; IPR004176; Clp_N.
DR InterPro; IPR001943; UvrB/C.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF02861; Clp_N; 2.
DR Pfam; PF02151; UVR; 1.
DR PRINTS; SM00382; AAA; 2.
DR PROSITE; PS00870; CLPAB_1; 1.
DR PROSITE; PS00871; CLPAB_2; 1.
DR PROSITE; PS0151; UVR; 1.
DR ATP-binding; Complete proteome; Protease.
SQ SEQUENCE    821 AA; 91174 MW; D0A3FF031346533F CRC64;

Query Match      81.4%; Score 35; DB 2; Length 821;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KIMDOVQQA 9
       |:|||||
Db      263 KIMDEIRQA 271

RESULT 7
Q8YST5
ID Q8YST5      PRELIMINARY;      PRT;      839 AA.
AC Q8YST5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Endopeptidase Clp ATP-binding chain.
GN Name=clpB; OrderedLocuNames=alr2999;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;

```

```

RESULT 6
Q8DML7
ID Q8DML7      PRELIMINARY;      PRT;      824 AA.
AC Q8DML7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ATP-dependent Clp protease regulatory subunit.
GN Name=clpC; OrderedLocuNames=t110307;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimp S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005369; BAC07860.1; -.
DR HSSP; P03815; 1JBK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR GO; GO:0019538; P:protein metabolism; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR001270; Chaprinin_clpA/B.
DR InterPro; IPR004176; Clp_N.
DR InterPro; IPR001943; UvrB/C.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF02861; Clp_N; 2.
DR Pfam; PF02151; UVR; 1.
DR PRINTS; SM00382; AAA; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00870; CLPAB_1; 1.
DR PROSITE; PS00871; CLPAB_2; 1.
DR PROSITE; PS0151; UVR; 1.
DR ATP-binding; Complete proteome; Protease.
SQ SEQUENCE    824 AA; 91507 MW; ECF7FFAA481A57CF CRC64;

Query Match      81.4%; Score 35; DB 2; Length 824;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KIMDOVQQA 9
       |:|||||
Db      262 KIMDEIRQA 270

RESULT 7
Q8YST5
ID Q8YST5      PRELIMINARY;      PRT;      839 AA.
AC Q8YST5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Endopeptidase Clp ATP-binding chain.
GN Name=clpB; OrderedLocuNames=alr2999;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;

```

```

RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003591; BAB74698.1; -.
DR PIR; AH2180; AH2180.
DR HSP; P03815; IJKB.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR GO; GO:0019538; P:protein metabolism; IEA.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF02861; CLP_N; 2.
DR PRINTS; PR00300; CLP_PPTEASEA.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00870; CLPAB_1; 1.
DR PROSITE; PS00871; CLPAB_2; 1.
DR PROSITE; PS50151; UVR; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 839 AA; 93140 MW; 948047142F9DB350 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 839;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9
Db 278 KIMDEIRQA 286

RESULT 8
EG51 XENLA STANDARD; PRT; 1060 AA.
AC P28025;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Kinesin-related motor protein Egs 1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCB1_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Egg;
RX MEDLINE=91246212; PubMed=1710028;
RA le Guellec R., Paris J., Couturier A., Roghi C., Philippe M.;
RT "Cloning by differential screening of a Xenopus cDNA that encodes a
RL Mol. Cell. Biol. 11:3395-3398(1991).
CC -!- FUNCTION: Plus end-directed motor protein required for
CC establishing a bipolar spindle. Associates with both interphase
CC and spindle microtubules. May be involved in nuclear divisions
CC taking place during the development of unfertilized eggs.
CC -!- SUBUNIT: Heterotetramer of two heavy and two light chains.
CC -!- SUBCELLULAR LOCATION: Concentrated around the polar ends of both
CC meiotic and mitotic spindles.
CC -!- TISSUE SPECIFICITY: Highly expressed in unfertilized eggs,
CC especially in the germinal vesicle and in the radial yolk-poor
CC channels. Also present in testis.
CC -!- DEVELOPMENTAL STAGE: Egg maturation. Is no longer synthesized soon
CC after fertilization.
CC -!- PTM: Phosphorylation of Thr-930 during mitosis controls the
CC association of Egs with the spindle apparatus (By similarity).

```

```

CC CC -!- SIMILARITY: Belongs to the kinesin-like protein family. BimC
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X54002; CAA37950.1; -.
CC PIR; A40264; A40264.
CC HSP; P52732; I1I6.
CC InterPro; IPR001752; kinesin_motor.
CC Pfam; PF00245; Kinesin; 1.
CC PRINTS; PR00380; KINESINHEAVY.
CC SMART; SM00129; KISC; 1.
CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
CC PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Coiled coil; Microtubule; Mitosis; Motor protein;
KW Phosphorylation.
KW DOMAIN 9 356 Kinesin-motor (By similarity).
FT DOMAIN 357 479 Coiled coil (Potential).
FT DOMAIN 688 729 Coiled coil (Potential).
FT DOMAIN 868 895 Coiled coil (Potential).
FT DOMAIN 1033 1060 Coiled coil (Potential).
FT NP_BIND 98 105 ATP (By similarity).
FT MOD_RES 930 930 Phosphothreonine (by CDC2) (By
FT similarity).
SQ SEQUENCE 1060 AA; 119332 MW; 201E7EB6962B3492 CRC64;

Query Match 81.4%; Score 35; DB 1; Length 1060;
Best Local Similarity 77.8%; Pred. No. 3.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9
Db 1040 KIMDEVQQA 1048

RESULT 9
Q632T3 PRELIMINARY; PRT; 1067 AA.
ID Q632T3
AC Q632T3, 2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCB1_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang S.I., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

```

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Griewood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Klein S., Gerhard D.S.;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC082827; AA082827.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 1067 AA; 119927 MW; B27C04D2196590AD CRC64;  
Query Match 81.4%; Score 35; DB 2; Length 1067;  
Best Local Similarity 77.8%; Pred. No. 3.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 KIMDQVQQA 9  
Db 1047 KIMDEVDQA 1055  
RESULT 10  
Q7NES9  
ID Q7NES9 PRELIMINARY; PRT; 80 AA.  
AC Q7NES9;  
DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Gsl3799 protein.  
GN OrderedLocusNames=gsl3799;  
OS Gloeobacter violaceus.  
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.  
OX NCBI\_TaxID=33072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PCC 7421;  
RX MEDLINE=22977040; PubMed=14621292;  
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,  
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,  
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a  
RT cyanobacterium that lacks thylakoids.";  
RL DNA Res. 10:137-145 (2003)  
DR EMBL; AP006581; BAC91740.1; -;  
DR InterPro; IPR006442; Pfd.  
DR TIGRFAMs; TIGR01552; pfd\_fam; 1.  
KW Complete proteome.  
SQ SEQUENCE 80 AA; 8685 MW; 861F05PE2D478B8E CRC64;  
Query Match 79.1%; Score 34; DB 2; Length 80;  
Best Local Similarity 75.0%; Pred. No. 40;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 KIMDQVQQA 8  
Db 19 KIMDQVQQA 26  
RESULT 11  
NHPX CAEEL  
ID \_NHPX\_CAEEL STANDARD; PRT; 128 AA.

AC Q21568;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE NHP2/L7AE family protein YEL026W homolog.  
GN ORFNames=M28.5;  
OS Caenorhabditis elegans.  
OC Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RG The C. elegans sequencing consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology.";  
RL Science 282:1012-1018 (1998).  
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -!- SIMILARITY: Belongs to the ribosomal protein L7AE family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z49911; CAA90127.1; -;  
DR PIR; T23808; T23808.  
DR HSSP; P55769; IE7K.  
DR WormBase; WBGene00010896; M28.5.  
DR InterPro; IPR004037; Ribosomal\_L7AE.  
DR Pfam; PF01248; Ribosomal\_L7AE; 1.  
DR PRINTS; PR00881; L7ARSEFAMILY.  
DR PRODOM; PD004495; Ribosomal\_L7AE; 1.  
DR PROSITE; PS01082; RIBOSOMAL\_L7AE; 1.  
KW Nuclear protein.  
SQ SEQUENCE 128 AA; 14001 MW; 445C4BF3504E45F0 CRC64;  
Query Match 79.1%; Score 34; DB 1; Length 128;  
Best Local Similarity 77.8%; Pred. No. 64;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 KIMDQVQQA 9  
Db 21 KIMDLVQQA 29  
RESULT 12  
Q9ZEG1  
ID Q9ZEG1 PRELIMINARY; PRT; 344 AA.  
AC Q9ZEG1;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ClpC protein (Fragment).  
GN NameclpC;  
OS Bacillus anthracis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Namy O., Mock M., Fouet A.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ224158; CAA11855.1; -;  
DR HSSP; P03815; 1JJBK.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0004518; F:nuclease activity; IEA.

GO: GO:0000166; F:nucleotide binding; IEA.  
GO: GO:0006289; P:nucleotide-excision repair; IEA.  
InterPro: IPR001593; AAA\_ATPase.  
InterPro: IPR003959; AAA\_ATPase\_cent.  
InterPro: IPR001270; Chaprinin\_c1pA/B.  
InterPro: IPR001943; UvrB/C.  
Pfam: PF00004; AAA; 1.  
Pfam: PF02151; UVR; 1.  
SMART: SM00382; AAA; 1.  
PROSITE: PS00870; CLPAB\_1; 1.  
PROSITE: PS50151; UVR; 1.  
ATP-binding. 1 1  
FT NON TER 344 344  
FT NON TER 344 344  
SQ SEQUENCE 344 AA; DBE73EFD03304C2B CRC64;  
  
Query Match 79.1%; Score 34; DB 2; Length 344;  
Best Local Similarity 55.6%; Pred. No. 1.7e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KIMDOVQQA 9  
|:|:|:|  
53 KVMDEIRQA 61  
  
Db  
  
RESULT 13  
SR54\_RICCN  
ID \_SR54\_RICCN STANDARD; PRT; 449 AA.  
AC Q92J55;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Signal recognition particle protein (Fifty-four homolog).  
DE Name=fff; OrderedLocusNames=RC0214;  
OS Rickettsia conorii.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
OC Rickettsiaceae; Rickettsiae; Rickettsia.  
NCBI\_TaxID=781;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Malish 7;  
RC MEDLINE=21442074; PubMed=11557893; DOI=10.1126/science.1061471;  
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,  
RA Sanson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,  
RA Raoult D.;  
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";  
RL Science 293:2093-2098 (2001).  
CC -!- FUNCTION: Necessary for efficient export of extra-cytoplasmic  
proteins. Binds to the signal sequence when it emerges from the  
ribosomes (By similarity).  
CC -!- SUBUNIT: Signal recognition particle consists of a 4.5 RNA  
molecule and protein fff (By similarity).  
CC -!- DOMAIN: The protein has a two domain structure: the G-domain binds  
GTP; the M-domain binds the RNA and also binds the signal  
sequence.  
CC -!- SIMILARITY: Belongs to the GTP-binding SRP family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC  
CC EMBL; AE008589; AAU02752.1; -  
DR PIR; F97726; F97726.  
DR HGSP; O07347; 1LS1.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR000897; SRP54.  
DR InterPro; IPR004125; SRP54\_SBP.  
DR InterPro; IPR004780; SRP\_sub.  
DR Pfam; PF000448; SRP54; 1.-

```

CC interact with mecA to negatively regulate comK.
CC -I- SIMILARITY: Belongs to the clpC/clpB family. ClpC subfamily.
CC CC -I- SIMILARITY: Contains 1 UVR domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; D26185; BAA05320.1; -;
DR ENBL; U02604; AAA19233.1; -;
DR ENBL; Z99104; CAB11862.1; -;
DR ENBL; X75930; CAA53534.1; -;
DR PIR; I40508; I40508.
DR HSSP; P03815; IJBK.
DR Subtilist; BG10148; clpC.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR001270; Chaprinin_clpA/B.
DR InterPro; IPR004176; Clp_N.
DR InterPro; IPR001943; UvrE/C.
DR Pfam; PF00004; AAA; 2.
DR Pfam; PF02861; Clp_N; 2.
DR Pfam; PF02151; UVR; 1.
DR PRINTS; PR00300; CLPPROTEASEA.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00870; CLPAB_1; 1.
DR PROSITE; PS00871; CLPAB_2; 1.
DR PROSITE; PS0151; UVR; 1.
KW ATP-binding; Chapterone; Competence; Complete proteome; Repeat;
KW Repressor; Transcription regulation.
FT DOMAIN 19 124 2 X 32 AA approximate repeats.
FT REPEAT 19 50 1.
FT REPEAT 93 124 2.
FT DOMAIN 163 410 1.
FT DOMAIN 417 452 UVR.
FT DOMAIN 471 662 II.
FT NP BIND 208 215 ATP (Potential).
FT NP BIND 545 552 ATP (Potential).
SQ SEQUENCE 810 AA; 90118 MW; 9592231886C2E870 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 810;
Best Local Similarity 55.6%; Pred. No. 3.9e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIMDQVQQA 9
Db 263 KWVDEIRQA 271
|:|:|:|:|
|:|:|:|:|

Search completed: November 18, 2005, 01:07:49
Job time : 46,9677 secs

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 00:48:52 ; Search time 13.6452 Seconds  
(Without alignments)  
49.237 Million cell updates/sec

Title: US-10-006-177-1

Perfect score: 43

Sequence: 1 KIMDOVQQA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	2842	1	US-07-741-940-7
2	43	100.0	2842	1	US-08-289-548A-7
3	43	100.0	2842	1	US-08-452-654-7
4	43	100.0	2842	1	US-08-449-731-7
5	43	100.0	2843	1	US-07-741-940-2
6	43	100.0	2843	1	US-08-289-548A-2
7	43	100.0	2843	1	US-08-452-654-2
8	43	100.0	2843	1	US-08-452-655B-2
9	43	100.0	2843	1	US-08-452-655B-7
10	43	100.0	2843	2	US-08-370-235A-2
11	43	100.0	2843	3	US-08-450-582-2
12	43	100.0	2843	3	US-08-450-582-7
13	43	100.0	2843	4	US-08-449-731-2
14	43	100.0	2843	4	US-10-092-138A-30
15	43	100.0	2843	4	US-09-538-032-1007
16	43	100.0	2843	2	US-08-821-335A-7
17	43	100.0	2973	2	US-09-003-687A-7
18	43	100.0	2973	3	US-09-136-605-7
19	34	79.1	649	4	US-09-902-540-13928
20	34	79.1	672	3	US-09-040-843-4
21	34	79.1	672	3	US-09-621-855-4
22	34	79.1	866	3	US-09-040-843-2
23	34	79.1	866	3	US-09-621-855-2
24	32	74.4	516	4	US-09-248-796A-17736
25	31	72.1	275	4	US-09-543-681A-4425
26	31	72.1	434	2	US-08-989-925-1
27	31	72.1	434	4	US-09-489-039A-7886

#### ALIGNMENTS

##### RESULT 1

US-07-741-940-7

; Sequence 7, Application US/07741940

; Patent No. 5352775

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS

; APPLICANT: ANAND, RAKESH

; APPLICANT: CARLSON, MARY

; APPLICANT: GRODEN, JOANNA

; APPLICANT: HEDGE, PHILIP J.

; APPLICANT: JOSLYN, GEOFF

; APPLICANT: KINZLER, KENNETH

; APPLICANT: MARKHAM, ALEXANDER F.

; APPLICANT: NAKAMURA, YUSUKE

; APPLICANT: THLIVERIS, ANDREW

; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS

; NUMBER OF SEQUENCES: 94

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner, Birch, McKie & Beckett

; STREET: 1001 G Street, NW

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20001-4598

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/741,940

; FILING DATE: 19920109

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Kagan, Sarah A.

; REGISTRATION NUMBER: 32,141

; REFERENCE/DOCKET NUMBER: 1107.035574

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-508-9100

; TELEFAX: 202-508-9299

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2842 amino acids

; TYPE: AMINO ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

Sequence 22, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 143, App  
Sequence 126, App  
Sequence 4, Appl  
Sequence 1418, Ap  
Sequence 4406, Ap  
Sequence 29210, A  
Sequence 9430, Ap  
Sequence 32144, A  
Sequence 10068, A  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 196, App  
Sequence 71, Appl  
Sequence 57437, A

28 31 72.1 452 4 US-09-010-147B-22  
29 31 72.1 491 2 US-08-942-819-2  
30 31 72.1 491 4 US-09-522-955A-2  
31 31 72.1 498 4 US-09-198-452A-143  
32 31 72.1 498 4 US-09-438-185A-126  
33 31 72.1 686 2 US-08-849-480A-4  
34 31 72.1 263 4 US-09-710-279-1418  
35 30 69.8 274 3 US-09-134-001C-4406  
36 30 69.8 563 4 US-09-252-991A-29210  
37 30 69.8 611 4 US-09-949-016-9430  
38 30 69.8 683 4 US-09-252-991A-32144  
39 30 69.8 751 4 US-09-902-540-10068  
40 30 69.8 1066 3 US-09-541-782-8  
41 30 69.8 1066 4 US-09-723-820-8  
42 30 69.8 1066 4 US-10-270-085-8  
43 29 67.4 125 4 US-09-732-210-196  
44 29 67.4 132 2 US-08-649-991-71  
45 29 67.4 156 4 US-09-270-767-57437

IMMEDIATE SOURCE:  
CLONE: APC  
US-07-741-940-7

Query Match 100.0%; Score 43; DB 1; Length 2842;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMQVQQA 9

Db 1744 KIMQVQQA 1752

RESULT 2

US-08-289-548A-7  
Sequence 7, Application US/08289548A  
Patent No. 5648212  
GENERAL INFORMATION:  
APPLICANT: ALBERTSEN, HANS  
APPLICANT: ANAND, RAKESH  
APPLICANT: CARLSON, MARY  
APPLICANT: GRODEN, JOANNA  
APPLICANT: HEDGE, PHILIP J.  
APPLICANT: JOSLYN, GEOFF  
APPLICANT: KINZLER, KENNETH  
APPLICANT: MARKHAM, ALEXANDER F.  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: THLIVERIS, ANDREW  
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Allegretti, LTD  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/289,548A  
FILING DATE: 12-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.46943  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2842 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: APC  
US-08-289-548A-7

Query Match 100.0%; Score 43; DB 1; Length 2842;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMQVQQA 9

|||||

Db 1744 KIMQVQQA 1752

RESULT 3

US-08-452-654-7  
Sequence 7, Application US/08452654  
Patent No. 5691454  
GENERAL INFORMATION:  
APPLICANT: ALBERTSEN, HANS  
APPLICANT: ANAND, RAKESH  
APPLICANT: CARLSON, MARY  
APPLICANT: GRODEN, JOANNA  
APPLICANT: HEDGE, PHILIP J.  
APPLICANT: JOSLYN, GEOFF  
APPLICANT: KINZLER, KENNETH  
APPLICANT: MARKHAM, ALEXANDER F.  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: THLIVERIS, ANDREW  
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,654  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.035574  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2842 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: APC  
US-08-452-654-7

Query Match 100.0%; Score 43; DB 1; Length 2842;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMQVQQA 9

|||||

Db 1744 KIMQVQQA 1752

RESULT 4

US-08-449-731-7  
Sequence 7, Application US/08449731  
Patent No. 6413727



```

; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; ANAND, RAKESH
; CARLSON, MARY
; GRODEN, JOANNA
; HEDGE, PHILIP J.
; JOSLYN, GEOFF
; KINZLER, KENNETH
; MARKHAM, ALEXANDER F.
; NAKAMURA, YUSUKE
; THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; GENE IN COLORECTAL CANCER IN HUMANS
;
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,731
; FILING DATE: 25-May-1995
; CLASSIFICATION: <unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/289,548
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.46943
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
;
; INFORMATION FOR SEQ ID NO: 7:
; CLONE: APC
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-449-731-7

Query Match 100.0%; Score 43; DB 4; Length 2842;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9
Db 1744 KIMDQVQQA 1752

RESULT 5
US-07-741-940-2
; Sequence 2, Application US/07741940
; Patent No. 5352775
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; ANAND, RAKESH
; CARLSON, MARY
; GRODEN, JOANNA
; HEDGE, PHILIP J.
; JOSLYN, GEOFF
; KINZLER, KENNETH
; MARKHAM, ALEXANDER F.
; NAKAMURA, YUSUKE
; THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; GENE IN COLORECTAL CANCER IN HUMANS
;
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598

```

```

; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; GENE IN COLORECTAL CANCER IN HUMANS
;
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/741,940
; FILING DATE: 19920109
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.035574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-741-940-2

Query Match 100.0%; Score 43; DB 1; Length 2843;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9
Db 1745 KIMDQVQQA 1753

RESULT 6
US-08-289-548A-2
; Sequence 2, Application US/08289548A
; Patent No. 5648212
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; ANAND, RAKESH
; CARLSON, MARY
; GRODEN, JOANNA
; HEDGE, PHILIP J.
; JOSLYN, GEOFF
; KINZLER, KENNETH
; MARKHAM, ALEXANDER F.
; NAKAMURA, YUSUKE
; THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; GENE IN COLORECTAL CANCER IN HUMANS
;
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598

```

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/289,548A  
; FILING DATE: 12-AUG-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.46943  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-289-548A-2

Query Match 100.0%; Score 43; DB 1; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
Db 1745 KIMDQVQQA 1753

RESULT 7  
US-08-654-2  
; Sequence 2, Application US/08452654  
; Patent No. 5691454  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 94  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner, Birch, McKie & Beckett  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/452,654  
; FILING DATE: 25-MAY-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/741,940  
; FILING DATE: 08-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141

; REFERENCE/DOCKET NUMBER: 1107.035574  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-452-654-2

Query Match 100.0%; Score 43; DB 1; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
Db 1745 KIMDQVQQA 1753

RESULT 8  
US-08-452-655B-2  
; Sequence 2, Application US/08452655B  
; Patent No. 5783666  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/452,655B  
; FILING DATE: 25-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/289,548  
; FILING DATE: 12-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/741,940  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.49964  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-452-655B-2

Query Match 100.0%; Score 43; DB 1; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIMDQVQQA 9  
DB 1745 KIMDQVQQA 1753

RESULT 9

US-08-452-655B-7  
; Sequence 7, Application US/08452655B  
; Patent No. 5783666  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/452,655B  
; FILING DATE: 25-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/289,548  
; FILING DATE: 12-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/741,940  
; FILING DATE: 08-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.49964  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ANTI-SENSE: NO

Query Match 100.0%; Score 43; DB 1; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIMDQVQQA 9  
DB 1745 KIMDQVQQA 1753

RESULT 10

US-08-370-235A-2  
; Sequence 2, Application US/08370235A  
; Patent No. 5910418  
; GENERAL INFORMATION:  
; APPLICANT: VOGELSTEIN, BERT  
; APPLICANT: KINZLER, KENNETH W.  
; APPLICANT: HILL, DAVID E.  
; APPLICANT: JOHNSON, KAREN A.  
; TITLE OF INVENTION: ANTIBODIES AND ASSAYS FOR DETERMINING  
; TITLE OF INVENTION: MUTATIONS IN THE APC GENE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BANNER & WITCOFF, LTD.  
; STREET: 1001 G STREET, N.W.  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: US  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370,235A  
; FILING DATE: 01-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KAGAN, SARAH A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 01107.48688  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202 508 9100  
; TELEFAX: 202 508 9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-370-235A-2

Query Match 100.0%; Score 43; DB 2; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIMDQVQQA 9  
DB 1745 KIMDQVQQA 1753

RESULT 11

US-08-450-582-2  
; Sequence 2, Application US/08450582  
; Patent No. 6114124  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW

;; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
;; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
;; NUMBER OF SEQUENCES: 102  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Banner & Witcoff, Ltd.  
;; STREET: 1001 G Street, NW  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20001-4598  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/450,582  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/452,655  
;; FILING DATE: 25-MAY-1995  
;; APPLICATION NUMBER: US 08/289,548  
;; FILING DATE: 12-AUG-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/741,940  
;; FILING DATE: 08-AUG-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kagan, Sarah A.  
;; REGISTRATION NUMBER: 32,141  
;; REFERENCE/DOCKET NUMBER: 1107.49964  
;; TELEPHONE: 202-508-9100  
;; TELEFAX: 202-508-9299  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2843 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; ANTI-SENSE: NO  
US-08-450-582-2

Query Match 100.0%; Score 43; DB 3; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 5.2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Oy 1 KIMDQVQQA 9  
Db 1745 KIMDQVQQA 1753

RESULT 12  
US-08-450-582-7  
; Sequence 7, Application US/08450582  
; Patent No. 6114124  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, NW  
; CITY: Washington

;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20001-4598  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/450,582  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/452,655  
;; FILING DATE: 25-MAY-1995  
;; APPLICATION NUMBER: US 08/289,548  
;; FILING DATE: 12-AUG-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/741,940  
;; FILING DATE: 08-AUG-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kagan, Sarah A.  
;; REGISTRATION NUMBER: 32,141  
;; REFERENCE/DOCKET NUMBER: 1107.49964  
;; TELEPHONE: 202-508-9100  
;; TELEFAX: 202-508-9299  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2843 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: YES  
;; ANTI-SENSE: NO  
US-08-450-582-7

Query Match 100.0%; Score 43; DB 3; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 5.2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Oy 1 KIMDQVQQA 9  
Db 1745 KIMDQVQQA 1753

RESULT 13  
US-08-449-731-2  
; Sequence 2, Application US/08449731  
; Patent No. 6413727  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Allegretti, LTD  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/449,731
; APPLICATION DATE: 25-May-1995
; FILING DATE: 25-May-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/289,548
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.46943
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-449-731-2

Query Match          100.0%; Score 43; DB 4; Length 2843;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 KIMDQVQQA 9
Db      1745 KIMDQVQQA 1753
```

```
RESULT 14
US-10-092-138A-30
; Sequence 30, Application US/10092138A
; Patent No. 6743630
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
; FILE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: 6S823/JPM/PT
; CURRENT APPLICATION NUMBER: US/10/092.138A
; CURRENT FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: human
US-10-092-138A-30
```

```
Query Match          100.0%; Score 43; DB 4; Length 2843;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 KIMDQVQQA 9
Db      1745 KIMDQVQQA 1753
```

```
RESULT 15
US-09-538-092-1007
; Sequence 1007, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
```

```
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurafatSeqFormatter Version 0.9
; SEQ ID NO 1007
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P25054
US-09-538-092-1007
```

```
Query Match          100.0%; Score 43; DB 4; Length 2843;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 KIMDQVQQA 9
Db      1745 KIMDQVQQA 1753
```

```
Search completed: November 18, 2005, 09:35:52
Job time : 16.6452 secs
```

**This Page Blank (uspio)**

GenCore version 5.1.6

Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 01:08:37 ; Search time 44.3226 Seconds  
(without alignments)  
84.961 Million cell updates/sec

Title: US-10-006-177-1

Perfect score: 43

Sequence: 1 KIMDQVQQA 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	9	14	US-10-006-177-1
2	43	100.0	902	18	US-10-450-763-39694
3	43	100.0	912	9	US-09-987-482-2
4	43	100.0	2843	8	US-08-681-219-32
5	43	100.0	2843	9	US-09-987-482-1
6	43	100.0	2843	10	US-09-230-111C-30
7	43	100.0	2843	14	US-10-092-138-30
8	43	100.0	2843	15	US-10-392-113-21
9	43	100.0	2843	16	US-10-408-765A-1970
10	43	100.0	2843	16	US-10-920-403-30
11	43	100.0	2844	15	US-10-267-502-370

12	43	100.0	2845	15	US-10-267-502-372	Sequence 372, App
13	37	86.0	348	15	US-10-282-122A-64196	Sequence 64196, A
14	35	81.4	817	15	US-10-369-493-18772	Sequence 18772, A
15	35	81.4	821	15	US-10-369-493-2791	Sequence 2791, Ap
16	35	81.4	821	17	US-10-732-923-6941	Sequence 6941, Ap
17	35	81.4	824	17	US-10-732-923-7015	Sequence 7015, Ap
18	35	81.4	839	17	US-10-732-923-6928	Sequence 6928, Ap
19	34	79.1	177	9	US-09-812-350-5	Sequence 5, Appli
20	34	79.1	177	17	US-10-732-923-6732	Sequence 6732, Ap
21	34	79.1	344	17	US-10-732-923-6731	Sequence 6731, Ap
22	34	79.1	809	15	US-10-369-493-16530	Sequence 16530, A
23	34	79.1	810	9	US-09-815-242-5586	Sequence 5586, Ap
24	34	79.1	810	15	US-10-369-493-23050	Sequence 23050, A
25	34	79.1	810	17	US-10-732-923-7056	Sequence 7056, Ap
26	34	79.1	811	15	US-10-282-122A-45773	Sequence 45773, A
27	34	79.1	811	17	US-10-732-923-7004	Sequence 7004, Ap
28	34	79.1	811	17	US-10-732-923-7067	Sequence 7067, Ap
29	34	79.1	813	15	US-10-369-493-17101	Sequence 17101, A
30	34	79.1	813	17	US-10-732-923-7201	Sequence 7201, Ap
31	34	79.1	818	9	US-09-815-242-12424	Sequence 12424, A
32	34	79.1	818	9	US-09-815-242-12789	Sequence 12789, A
33	34	79.1	818	15	US-10-282-122A-44278	Sequence 44278, A
34	34	79.1	818	17	US-10-732-923-6968	Sequence 6968, Ap
35	34	79.1	818	17	US-10-732-923-7010	Sequence 7010, Ap
36	34	79.1	820	15	US-10-282-122A-60729	Sequence 60729, A
37	34	79.1	820	17	US-10-732-923-6979	Sequence 6979, Ap
38	34	79.1	820	17	US-10-732-923-6982	Sequence 6982, Ap
39	34	79.1	825	17	US-10-732-923-6975	Sequence 6975, Ap
40	33	76.7	1443	16	US-10-437-963-114387	Sequence 114387, A
41	32	74.4	272	9	US-09-925-297-661	Sequence 661, App
42	32	74.4	392	14	US-10-181-071-5	Sequence 5, Appli
43	32	74.4	513	16	US-10-739-930-6710	Sequence 6710, Ap
44	32	74.4	648	14	US-10-032-585-7540	Sequence 7540, Ap
45	32	74.4	707	16	US-10-437-963-145806	Sequence 145806, A

## ALIGNMENTS

RESULT 1  
US-10-006-177-1  
; Sequence 1, Application US/10006177  
; Publication No. US20030165513A1  
; GENERAL INFORMATION:  
; APPLICANT: Ramakrishna, Venky  
; APPLICANT: Ross, Mark  
; APPLICANT: Philip, Ramila  
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treatm  
; FILE REFERENCE: 26747-35  
; CURRENT APPLICATION NUMBER: US/10/006.177  
; CURRENT FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: US/60/251,022  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: US/60/256,824  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Epitopic Peptide  
US-10-006-177-1

Query Match 100.0%; Score 43; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9

Db 1 KIMDQVQQA 9

RESULT 2  
 US-10-450-763-39694  
 ; Sequence 39694, Application US/10450763  
 ; Publication No. US20050196754A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
 ; FILE REFERENCE: 790CIP3/US  
 ; CURRENT APPLICATION NUMBER: US/10/450,763  
 ; CURRENT FILING DATE: 2003-06-11  
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: 09/540,217  
 ; PRIOR FILING DATE: 2000-03-31  
 ; PRIOR APPLICATION NUMBER: 09/649,167  
 ; PRIOR FILING DATE: 2000-08-23  
 ; NUMBER OF SEQ ID NOS: 60736  
 ; SOFTWARE: Custom  
 ; SEQ ID NO 39694  
 ; LENGTH: 902  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: DOMAIN  
 ; LOCATION: (73)..(107)  
 ; OTHER INFORMATION: Armadillo/beta-catenin-like repeat proteins domain identified  
 ; OTHER INFORMATION: by EMATRIX, accession number PF00514C, p-value=5.250e-17, raw  
 ; OTHER INFORMATION: score of 24.37  
 US-10-450-763-39694

Query Match 100.0%; Score 43; DB 18; Length 902;  
 Best Local Similarity 100.0%; Pred. No. 8.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
 |||||  
 Db 277 KIMDQVQQA 285

RESULT 3  
 US-09-987-482-2  
 ; Sequence 2, Application US/09987482  
 ; Publication No. US20020184656A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BHANDARI, POONAM  
 ; APPLICANT: SHASHIDHARA, L.S.  
 ; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF  
 ; FILE REFERENCE: 056859-0134  
 ; CURRENT APPLICATION NUMBER: US/09/987,482  
 ; CURRENT FILING DATE: 2002-03-21  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 912  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-987-482-2

Query Match 100.0%; Score 43; DB 9; Length 912;  
 Best Local Similarity 100.0%; Pred. No. 8.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
 |||||  
 Db 787 KIMDQVQQA 795

RESULT 4  
 US-08-681-219-32  
 ; Sequence 32, Application US/08681219

; Publication No. US20020058607A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Takaaki Sato and Junn Yanagisawa  
 ; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN  
 ; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/681,219  
 ; FILING DATE: 22-JUL-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 0575/48962/JPW/JKM  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0525  
 ; INFORMATION FOR SEQ ID NO: 32:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2843 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-681-219-32

Query Match 100.0%; Score 43; DB 8; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9  
 |||||  
 Db 1745 KIMDQVQQA 1753

RESULT 5  
 US-09-987-482-1  
 ; Sequence 1, Application US/09987482  
 ; Publication No. US20020184656A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BHANDARI, POONAM  
 ; APPLICANT: SHASHIDHARA, L.S.  
 ; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF  
 ; FILE REFERENCE: 056859-0134  
 ; CURRENT APPLICATION NUMBER: US/09/987,482  
 ; CURRENT FILING DATE: 2002-03-21  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 2843  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-987-482-1

Query Match 100.0%; Score 43; DB 9; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDQVQQA 9



```
Db      1745 KIMDQVQQA 1753
|||||||
RESULT 6
US-09-230-111C-30
; Sequence 30, Application US/09230111C
; Publication No. US20030203414A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; APPLICANT: Yanagisawa, Junn
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)
; TITLE OF INVENTION: DOMAIN AND USES THEREOF
; FILE REFERENCE: 48962-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,111C
; CURRENT FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: human
US-09-230-111C-30

Query Match      100.0%; Score 43; DB 10; Length 2843;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KIMDQVQQA 9
Db      1745 KIMDQVQQA 1753
|||||||

RESULT 7
US-10-092-138-30
; Sequence 30, Application US/10092138
; Publication No. US20030170723A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: 65823/JPW/PT
; CURRENT APPLICATION NUMBER: US/10/092,138
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: human
US-10-092-138-30

Query Match      100.0%; Score 43; DB 14; Length 2843;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KIMDQVQQA 9
Db      1745 KIMDQVQQA 1753
|||||||

RESULT 8
US-10-392-113-21
; Sequence 21, Application US/10392113
; Publication No. US20030224993A1
; GENERAL INFORMATION:
; APPLICANT: Land, Hartmut
; APPLICANT: Deleu, Laurent
; TITLE OF INVENTION: COMPOSITIONS THAT INHIBIT PROLIFERATION
; TITLE OF INVENTION: OF CANCER CELLS
; FILE REFERENCE: 21108.0005U3
; CURRENT APPLICATION NUMBER: US/10/392,113
```

```
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/365,078
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: PCT/US01/32127
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/239,705
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: Synthetic Construct
US-10-392-113-21

Query Match      100.0%; Score 43; DB 15; Length 2843;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KIMDQVQQA 9
Db      1745 KIMDQVQQA 1753
|||||||

RESULT 9
US-10-408-765A-1970
; Sequence 1970, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1970
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1970

Query Match      100.0%; Score 43; DB 16; Length 2843;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KIMDQVQQA 9
Db      1745 KIMDQVQQA 1753
|||||||

RESULT 10
US-10-820-403-30
; Sequence 30, Application US/10820403
; Publication No. US20040229287A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: 65823/JPW/PT
; CURRENT APPLICATION NUMBER: US/10/820,403
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: US/10/092,138
```

```

; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 2843
; TYPE: prt
; ORGANISM: human
US-10-820-403-30

```

```
Query Match      100.0%; Score 43; DB 16; Length 2843;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 KIMDQVQQA 9  
1745 KIMDQVQQA 1753

```

RESULT 11
US-10-267-502-370
; Sequence 370, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 370
; LENGTH: 2844
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-370

```

```
Query Match      100.0%; Score 43; DB 15; Length 2844;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 KIMDQVQQA 9  
Dp 1746 KIMDOVQQA 1754

```

RESULT 12
US-10-267-502-372
; Sequence 372, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSP-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 372
; LENGTH: 2845
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-267-502-372

```

```
Query Match      100.0%; Score 43; DB 15; Length 2845;
Best Local Similarity 100.0%; Pred. No. 29;
Matches      9: Conservative      0: Mismatches      0: Indels      0: Gaps      0:
```

Qy 1 KIMDQVQQA 9  
|||  
Db 1743 KIMDOVOOA 1751

RESULT 13  
US-10-282-122A-64196  
: Sequence 64196, Application US/10282122A  
: Publication No. US20040029129A1  
: GENERAL INFORMATION:  
: APPLICANT: Wang, Liangsu  
: APPLICANT: Zamudio, Carlos  
: APPLICANT: Malone, Cheryl  
: APPLICANT: Haselbeck, Robert  
: APPLICANT: Olsen, Kari  
: APPLICANT: Zyskind, Judith  
: APPLICANT: Wall, Daniel  
: APPLICANT: Trawick, John  
: APPLICANT: Carr, Grant  
: APPLICANT: Yamamoto, Robert  
: APPLICANT: Forsyth, R.  
: APPLICANT: Xu, H.

```

1  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
2
3  FILE REFERENCE: ELITRA.034A
4
5  CURRENT APPLICATION NUMBER: US/10/282,122A
6
7  CURRENT FILING DATE: 2003-02-20
8
9  PRIOR APPLICATION NUMBER: 60/191,078
10
11 PRIOR FILING DATE: 2000-03-21
12
13 PRIOR APPLICATION NUMBER: 60/206,848
14
15 PRIOR FILING DATE: 2000-05-23
16
17 PRIOR APPLICATION NUMBER: 60/207,727
18
19 PRIOR FILING DATE: 2000-05-26
20
21 PRIOR APPLICATION NUMBER: 60/230,335
22
23 PRIOR FILING DATE: 2000-09-06
24
25 PRIOR APPLICATION NUMBER: 60/230,347
26
27 PRIOR FILING DATE: 2000-09-09
28
29 PRIOR APPLICATION NUMBER: 60/242,578
30
31 PRIOR FILING DATE: 2000-10-23
32
33 PRIOR APPLICATION NUMBER: 60/253,625
34
35 PRIOR FILING DATE: 2000-11-27
36
37 PRIOR APPLICATION NUMBER: 60/257,931
38
39 PRIOR FILING DATE: 2000-12-22
40
41 PRIOR APPLICATION NUMBER: 60/267,636
42
43 PRIOR FILING DATE: 2001-02-09
44
45 PRIOR APPLICATION NUMBER: 60/269,308
46
47 PRIOR FILING DATE: 2001-02-16
48
49 Remaining Prior Application data removed - See File Wrapper or PALM.
50
51 NUMBER OF SEQ ID NOS: 78614
52
53 SOFTWARE: PatentIn version 3.1
54
55 SEQ ID NO 64196
56
57 LENGTH: 348
58
59 TYPE: PRT
60
61 ORGANISM: Mycoplasma pneumoniae
62
63 US-10-282-122A-64196

```

Query Match	86.0%;	Score 37;	DB 15;	Length 348;
Best Local Similarity	77.8%;	Pred. No. 46;		
Matches	7:	Conservative	2:	Mismatches
			0:	Indels
			0:	Gaps

QY 1 KIMDQVQQA 9  
|:|||||  
pB 16 KLYDPOVOOA 24

```

RESULT 14
US-10-369-493-18772
; Sequence 18772, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICR
; TITLE OF INVENTION: PLANTS WITH IMPRO
; FILE REFERENCE: 38-10(52052)B

```

; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 18772  
; LENGTH: 817  
; TYPE: PRT  
; ORGANISM: Anabaena PCC7120  
US-10-369-493-18772

Query Match 81.4%; Score 35; DB 15; Length 817;  
Best Local Similarity 66.7%; Pred. NO. 2.9e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDOVQQA 9  
Db 262 KIMDEIRQA 270

RESULT 15  
US-10-369-493-2791  
; Sequence 2791, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2791  
; LENGTH: 821  
; TYPE: PRT  
; ORGANISM: Synecocystis sp.  
US-10-369-493-2791

Query Match 81.4%; Score 35; DB 15; Length 821;  
Best Local Similarity 66.7%; Pred. NO. 2.9e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIMDOVQQA 9  
Db 263 KIMDEIRQA 271

Search completed: November 18, 2005, 10:44:27  
Job time : 47.3226 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:03:06 ; Search time 54.1936 Seconds  
(without alignments)  
71.366 Million cell updates/sec

Title: US-10-006-177-2

Perfect score: 53

Sequence: 1 RLQEDPPAGV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	100.0	10	5	ABG32302 HLA-A2 as
2	53	100.0	57	3	AAG02226 Human sec
3	53	100.0	122	5	ABP68915 Human pol
4	53	100.0	122	6	ABR43254 Human PMM
5	53	100.0	137	5	ABP68916 Human pol
6	49	92.5	81	3	AAG19027 Zea mays
7	49	92.5	90	3	AAG19026 Zea mays
8	49	92.5	141	7	ABM74252 DNA clone
9	49	92.5	143	3	AAG10241 Arabidops
10	49	92.5	143	3	AAG35168 Zea mays
11	49	92.5	143	3	AAG32531 Arabidops
12	49	92.5	151	4	ABB58767 Drosophil
13	49	92.5	152	2	AAR67277 Ubiquitin
14	49	92.5	152	2	AAR79654 Human E2
15	49	92.5	152	2	AAY39967 Human rad
16	49	92.5	152	3	AAB03175 Human rad
17	49	92.5	152	3	AAY70096 Zea mays
18	49	92.5	152	3	AAY70095 Zea mays
19	49	92.5	152	3	AAG10240 Arabidops
20	49	92.5	152	3	AAG35167 Zea mays
21	49	92.5	152	3	AAG32530 Arabidops
22	49	92.5	152	7	ADD47216 Rat Prote
23	49	92.5	152	7	ADE57919 Rat Prote
24	49	92.5	152	7	ADE58495 Rat Prote
25	49	92.5	152	7	ADD47218 Human Pro

## ALIGNMENTS

### RESULT 1

ABG32302

ID ABG32302 standard; peptide; 10 AA.

XX AC ABG32302;

XX XX

DT 05-NOV-2002 (first entry)

XX XX

DE HLA-A2 associated immunogenic peptide from human HHR6A protein.

XX Human; immunogen; epitope; HLA-A1; human leukocyte antigen; CTL;

KW cytotoxic Y lymphocyte; cytostatic; cancer; colorectal carcinoma;

KW ovarian carcinoma; lung carcinoma; prostate carcinoma; vaccine; tumour;

KW HLA-2; passive immunotherapy; HHR6A; ubiquitin conjugating enzyme.

XX Homo sapiens.

OS

XX WO200246416-A2.

FN

XX 13-JUN-2002.

PD

PF 04-DEC-2001; 2001WO-US047290.

XX

PR 04-DEC-2000; 2000US-0251022P.

XX

PA 20-DEC-2000; 2000US-0256824P.

XX (ARGO-) ARGONEX INC.

XX Ramakrishna V, Ross M, Phillip R;

PI WPI; 2002-619021/66.

DR

XX New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte,

XX and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.

PT Claim 1; Page 50; 60pp; English.

XX

CC The invention relates to an immunogen comprising an isolated polypeptide

CC whose amino acid sequence comprises an epitopic peptide, does not include

CC MAGE 4 or MFG-E8 proteins, or consists of MAGE D protein or its

CC immunologically active fragment. Also included are a polynucleotide

CC encoding the immunogen or its complement, a vector comprising the

CC polynucleotide, a mammalian cell comprising the vector and expressing the

CC polynucleotide, a vaccine composition comprising the immunogen and an

CC antibody specific for the immunogen. The immunogen is useful for inducing

CC a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour

CC cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is

CC useful for inducing a CTL response when administered to a subject. A

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC mammalian cell that can express the immunogen, is useful for inducing a  
 CC CTL response in vitro that is specific for a tumour cell expressing HLA-1  
 CC or HLA-2. The immunogen or cell is useful for inducing CTL for treating a  
 CC subject with cancer (carcinoma, preferably colorectal carcinoma, ovarian  
 CC carcinoma, lung carcinoma and prostate carcinoma). The immunogen is also  
 CC useful for screening and diagnostic agents, for gene screening in  
 CC patients afflicted with cancer, for screening a sample for the presence  
 CC of CTLs that specifically recognise the corresponding epitopes, as a  
 CC diagnostic tool to evaluate the efficacy of the immunotherapeutic  
 CC treatments, to prepare class I MHC (major histocompatibility class)  
 CC tetramers which are utilised in conjunction with flow cytometry to  
 CC quantitate the frequency of peptide-specific CTL that are present in a  
 CC samples of lymphocytes from an individual, and for stimulating the  
 CC production of antibodies for use in passive immunotherapy, for use as  
 CC diagnostic reagents, and for use as reagents in other processes such as  
 CC affinity chromatography. The present sequence is an immunogenic epitope  
 CC of the invention derived from human HHR6A (ubiquitin conjugating enzyme  
 XX  
 XX  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 53; DB 5; Length 10;  
 Best Local Similarity 100.0%; Pred. NO. 0.026;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
 |||||  
 Db 1 RLQEDPPAGV 10

# RESULT 2

AAG02226 ID AAG02226 standard; protein; 57 AA.

XX AAG02226;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 6307.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX N-PSDB; AAC02232.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 13; SEQ ID NO 6307; 71pp + Sequence Listing; English.

XX The present sequence is a polypeptide encoded by one of a large number of  
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were  
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different  
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated  
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT  
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA  
 CC sequences derived from the 5' ends of mRNAs and even in those cases where  
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely

CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can  
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs  
 CC are also used in diagnostic, forensic, gene therapy and chromosome  
 CC mapping procedures. They are used to obtain upstream regulatory sequences  
 CC and to design expression and secretion vectors  
 XX  
 SQ Sequence 57 AA;

Query Match 100.0%; Score 53; DB 3; Length 57;  
 Best Local Similarity 100.0%; Pred. NO. 0.14;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
 |||||  
 Db 15 RLQEDPPAGV 24

# RESULT 3

ABP68915

XX ABP68915 standard; protein; 122 AA.

XX ABP68915;

XX 20-JAN-2003 (first entry)

XX Human polypeptide SEQ ID NO 962.

XX Human; genome mapping; gene therapy; food supplement; virus; fungus;  
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
 KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;  
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
 KW antiarthritic.

XX Homo sapiens.

XX WO200270539-A2.

XX 12-SEP-2002.

XX 05-MAR-2002; 2002WO-US005095.

XX 05-MAR-2001; 2001US-00799451.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
 PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2002-759812/82.  
 XX N-PSDB; ABZ11132.

XX New polynucleotides comprising sequences assembled from expressed  
 PT sequence tags (ESTs), useful for treating cell-proliferative, or platelet  
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or coagulation disorders.

XX Claim 9; SEQ ID NO 962; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a  
 CC nucleotide sequence selected from any of 948 sequences (ABZ11119-  
 CC ABZ12066) or their mature protein coding portion, active domain coding  
 CC protein or complementary sequences. The polynucleotides are useful for  
 CC identifying expressed genes or for physical mapping of human genome. The  
 CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
 CC markers, as a food supplement, for generating antibodies, in medical  
 CC imaging, screening and diagnostic assays and for treating cell-  
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,

CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
 CC arthritis, etc. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 122 AA;

Query Match 100.0%; Score 53; DB 5; Length 122;

Best Local Similarity 100.0%; Pred. No. 0.29; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0;

QY 1 RLQEDPPAGV 10

Db 15 RLQEDPPAGV 24

RESULT 4

ID ABR43254 standard; protein; 122 AA.

XX ABR43254;

DT 07-JUL-2003 (first entry)

DE Human PMMM-15 protein SEQ ID NO:15.

XX Human; protein modification and maintenance molecule; PMMM; cytostatic;  
 KW antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective; AIDS;  
 KW cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory; cancer;  
 KW thrombolytic; gene therapy; cell proliferative disorder; atherosclerosis;  
 KW neurological disorder; epilepsy; Huntington's disease; stroke; allergy;  
 KW immune disorder; inflammatory disorder; developmental disorder;  
 KW hypothyroidism; Cushing's syndrome; infection.

XX Homo sapiens.

XX WO2003025131-A2.

XX 27-MAR-2003.

XX 13-SEP-2002; 2002WO-US029221.

XX 14-SEP-2001; 2001US-0322196P.

XX 21-SEP-2001; 2001US-0324134P.

XX 05-OCT-2001; 2001US-0327233P.

XX 26-OCT-2001; 2001US-0346198P.

XX 02-NOV-2001; 2001US-0343980P.

XX 09-NOV-2001; 2001US-0348887P.

XX 16-NOV-2001; 2001US-0332423P.

XX 28-NOV-2001; 2001US-0334145P.

XX 06-DEC-2001; 2001US-0334229P.

XX 25-JAN-2002; 2002US-0351928P.

XX 21-MAR-2002; 2002US-0366837P.

XX (INCY-) INCYTE GENOMICS INC.

XX Sprague JW, Chawla NK, Warren BA, Tang YT, Elliott VS;

XX Marquis JP, Li JX, Griffin JA, Gietzen KJ, Yang J, Lu DAM;

XX Emerling BM, Duggan BM, Richardson TW, Lee SY, Ramkumar J, Becha SD;

XX Lehr-Mason PM, Swarnakar A, Tran UK, Kable AE, Hafalia AJA, Khare R;

XX WPI; 2003-354597/33.

XX N-PSDB; ACC59973.

XX New human protein modification and maintenance molecules (PMMM), useful  
 PT for diagnosing, treating and preventing diseases or conditions associated  
 PT with the aberrant PMMM expression e.g. cancer, AIDS, epilepsy, or  
 PT infections.

PS Claim 1; Page 225; 270pp; English.

XX

CC ACS59959 to ACC59989 encode the human protein modification and  
 CC maintenance molecule proteins given in ABR43240 to ABR43270, designated  
 CC PMMM-1 to PMMM-31 (1). (1) have cytostatic, antiarteriosclerotic,  
 CC anticonvulsant, neurotropic, neuroprotective, cerebroprotective, anti-HIV,  
 CC anti-allergic, anti-inflammatory and thymomimetic activities, and can be  
 CC used in gene therapy. The PMMM polypeptides and polynucleotides are  
 CC useful in diagnosing, treating and preventing diseases or conditions  
 CC associated with the decreased expression or overexpression of PMMM, such  
 CC as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g.  
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,  
 CC allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome)  
 CC disorders, or infections. They are also useful in assessing the effects  
 CC of exogenous compounds on the expression of nucleic acid and amino acid  
 CC sequences of PMMM. The PMMMs or their fragments are useful in screening  
 CC compounds for effectiveness as agonist or antagonist of the polypeptides,  
 CC or in altering the expression of the target polynucleotide and compounds  
 CC that specifically bind to or modulate the activity of the polypeptide.  
 CC The microarray is useful in monitoring or measuring protein-protein  
 CC interactions, drug-target interactions, and gene expression profiles

XX SQ Sequence 122 AA;

Query Match 100.0%; Score 53; DB 6; Length 122;

Best Local Similarity 100.0%; Pred. No. 0.29; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0;

QY 1 RLQEDPPAGV 10

Db 15 RLQEDPPAGV 24

RESULT 5

ID ABR68916 standard; protein; 137 AA.

XX ABR68916;

XX 20-JAN-2003 (first entry)

XX Human polypeptide SEQ ID NO 963.

XX Human; genome mapping; gene therapy; food supplement; virus; fungus;  
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
 KW arthritis; cytostatic; immunomodulator; neurotropic; neuroprotective;  
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
 KW antiarthritic.

XX Homo sapiens.

XX WO200270539-A2.

XX 12-SEP-2002.

XX 05-MAR-2002; 2002WO-US005095.

XX 05-MAR-2001; 2001US-00799451.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;

XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;

XX Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2002-759812/82.

XX N-PSDB; ABZ11133.

XX New polynucleotides comprising sequences assembled from expressed  
 PT sequence tags (ESTs), useful for treating cell-proliferative,  
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
 PT or coagulation disorders.

```
XX Claim 9; SEQ ID NO 963; 1012pp + Sequence Listing; English.
PS
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (AB211119-
CC AB212066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune disease (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 137 AA;

Query Match 100.0%; Score 53; DB 5; Length 137;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10
Db 15 RLQEDPPAGV 24
|||||

RESULT 6
AAG19027
ID AAG19027 standard; protein; 81 AA.
AC AAG19027;
XX
XX 17-OCT-2000 (first entry)
DT
XX
XX Zea mays protein fragment SEQ ID NO: 20669.
DE
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
XX Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131443P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132488P.
PR 11-MAY-1999; 99US-0132489P.
PR 14-MAY-1999; 99US-0132490P.
PR 14-MAY-1999; 99US-0132491P.
PR 14-MAY-1999; 99US-0132492P.
PR 14-MAY-1999; 99US-0132493P.
PR 14-MAY-1999; 99US-0132494P.
PR 14-MAY-1999; 99US-0132495P.
PR 14-MAY-1999; 99US-0132496P.
PR 14-MAY-1999; 99US-0132497P.
PR 14-MAY-1999; 99US-0132498P.
PR 14-MAY-1999; 99US-0132499P.
PR 14-MAY-1999; 99US-0132500P.
PR 14-MAY-1999; 99US-0132501P.
PR 14-MAY-1999; 99US-0132502P.
PR 14-MAY-1999; 99US-0132503P.
PR 14-MAY-1999; 99US-0132504P.
PR 14-MAY-1999; 99US-0132505P.
PR 14-MAY-1999; 99US-0132506P.
PR 14-MAY-1999; 99US-0132507P.
PR 14-MAY-1999; 99US-0132508P.
PR 14-MAY-1999; 99US-0132509P.
PR 14-MAY-1999; 99US-0132510P.
PR 14-MAY-1999; 99US-0132511P.
PR 14-MAY-1999; 99US-0132512P.
PR 14-MAY-1999; 99US-0132513P.
PR 14-MAY-1999; 99US-0132514P.
PR 14-MAY-1999; 99US-0132515P.
PR 14-MAY-1999; 99US-0132516P.
PR 14-MAY-1999; 99US-0132517P.
PR 14-MAY-1999; 99US-0132518P.
PR 14-MAY-1999; 99US-0132519P.
PR 14-MAY-1999; 99US-0132520P.
PR 14-MAY-1999; 99US-0132521P.
PR 14-MAY-1999; 99US-0132522P.
PR 14-MAY-1999; 99US-0132523P.
PR 14-MAY-1999; 99US-0132524P.
PR 14-MAY-1999; 99US-0132525P.
PR 14-MAY-1999; 99US-0132526P.
PR 14-MAY-1999; 99US-0132527P.
PR 14-MAY-1999; 99US-0132528P.
PR 14-MAY-1999; 99US-0132529P.
PR 14-MAY-1999; 99US-0132530P.
PR 14-MAY-1999; 99US-0132531P.
PR 14-MAY-1999; 99US-0132532P.
PR 14-MAY-1999; 99US-0132533P.
PR 14-MAY-1999; 99US-0132534P.
PR 14-MAY-1999; 99US-0132535P.
PR 14-MAY-1999; 99US-0132536P.
PR 14-MAY-1999; 99US-0132537P.
PR 14-MAY-1999; 99US-0132538P.
PR 14-MAY-1999; 99US-0132539P.
PR 14-MAY-1999; 99US-0132540P.
PR 14-MAY-1999; 99US-0132541P.
PR 14-MAY-1999; 99US-0132542P.
PR 14-MAY-1999; 99US-0132543P.
PR 14-MAY-1999; 99US-0132544P.
PR 14-MAY-1999; 99US-0132545P.
PR 14-MAY-1999; 99US-0132546P.
PR 14-MAY-1999; 99US-0132547P.
PR 14-MAY-1999; 99US-0132548P.
PR 14-MAY-1999; 99US-0132549P.
PR 14-MAY-1999; 99US-0132550P.
PR 14-MAY-1999; 99US-0132551P.
PR 14-MAY-1999; 99US-0132552P.
PR 14-MAY-1999; 99US-0132553P.
PR 14-MAY-1999; 99US-0132554P.
PR 14-MAY-1999; 99US-0132555P.
PR 14-MAY-1999; 99US-0132556P.
PR 14-MAY-1999; 99US-0132557P.
PR 14-MAY-1999; 99US-0132558P.
PR 14-MAY-1999; 99US-0132559P.
PR 14-MAY-1999; 99US-0132560P.
PR 14-MAY-1999; 99US-0132561P.
PR 14-MAY-1999; 99US-0132562P.
PR 14-MAY-1999; 99US-0132563P.
PR 14-MAY-1999; 99US-0132564P.
PR 14-MAY-1999; 99US-0132565P.
PR 14-MAY-1999; 99US-0132566P.
PR 14-MAY-1999; 99US-0132567P.
PR 14-MAY-1999; 99US-0132568P.
PR 14-MAY-1999; 99US-0132569P.
PR 14-MAY-1999; 99US-0132570P.
PR 14-MAY-1999; 99US-0132571P.
PR 14-MAY-1999; 99US-0132572P.
PR 14-MAY-1999; 99US-0132573P.
PR 14-MAY-1999; 99US-0132574P.
PR 14-MAY-1999; 99US-0132575P.
PR 14-MAY-1999; 99US-0132576P.
PR 14-MAY-1999; 99US-0132577P.
PR 14-MAY-1999; 99US-0132578P.
PR 14-MAY-1999; 99US-0132579P.
PR 14-MAY-1999; 99US-0132580P.
PR 14-MAY-1999; 99US-0132581P.
PR 14-MAY-1999; 99US-0132582P.
PR 14-MAY-1999; 99US-0132583P.
PR 14-MAY-1999; 99US-0132584P.
PR 14-MAY-1999; 99US-0132585P.
PR 14-MAY-1999; 99US-0132586P.
PR 14-MAY-1999; 99US-0132587P.
PR 14-MAY-1999; 99US-0132588P.
PR 14-MAY-1999; 99US-0132589P.
PR 14-MAY-1999; 99US-0132590P.
PR 14-MAY-1999; 99US-0132591P.
PR 14-MAY-1999; 99US-0132592P.
PR 14-MAY-1999; 99US-0132593P.
PR 14-MAY-1999; 99US-0132594P.
PR 14-MAY-1999; 99US-0132595P.
PR 14-MAY-1999; 99US-0132596P.
PR 14-MAY-1999; 99US-0132597P.
PR 14-MAY-1999; 99US-0132598P.
PR 14-MAY-1999; 99US-0132599P.
PR 14-MAY-1999; 99US-0132600P.
PR 14-MAY-1999; 99US-0132601P.
PR 14-MAY-1999; 99US-0132602P.
PR 14-MAY-1999; 99US-0132603P.
PR 14-MAY-1999; 99US-0132604P.
PR 14-MAY-1999; 99US-0132605P.
PR 14-MAY-1999; 99US-0132606P.
PR 14-MAY-1999; 99US-0132607P.
PR 14-MAY-1999; 99US-0132608P.
PR 14-MAY-1999; 99US-0132609P.
PR 14-MAY-1999; 99US-0132610P.
PR 14-MAY-1999; 99US-0132611P.
PR 14-MAY-1999; 99US-0132612P.
PR 14-MAY-1999; 99US-0132613P.
PR 14-MAY-1999; 99US-0132614P.
PR 14-MAY-1999; 99US-0132615P.
PR 14-MAY-1999; 99US-0132616P.
PR 14-MAY-1999; 99US-0132617P.
PR 14-MAY-1999; 99US-0132618P.
PR 14-MAY-1999; 99US-0132619P.
PR 14-MAY-1999; 99US-0132620P.
PR 14-MAY-1999; 99US-0132621P.
PR 14-MAY-1999; 99US-0132622P.
PR 14-MAY-1999; 99US-0132623P.
PR 14-MAY-1999; 99US-0132624P.
PR 14-MAY-1999; 99US-0132625P.
PR 14-MAY-1999; 99US-0132626P.
PR 14-MAY-1999; 99US-0132627P.
PR 14-MAY-1999; 99US-0132628P.
PR 14-MAY-1999; 99US-0132629P.
PR 14-MAY-1999; 99US-0132630P.
PR 14-MAY-1999; 99US-0132631P.
PR 14-MAY-1999; 99US-0132632P.
PR 14-MAY-1999; 99US-0132633P.
PR 14-MAY-1999; 99US-0132634P.
PR 14-MAY-1999; 99US-0132635P.
PR 14-MAY-1999; 99US-0132636P.
PR 14-MAY-1999; 99US-0132637P.
PR 14-MAY-1999; 99US-0132638P.
PR 14-MAY-1999; 99US-0132639P.
PR 14-MAY-1999; 99US-0132640P.
PR 14-MAY-1999; 99US-0132641P.
PR 14-MAY-1999; 99US-0132642P.
PR 14-MAY-1999; 99US-0132643P.
PR 14-MAY-1999; 99US-0132644P.
PR 14-MAY-1999; 99US-0132645P.
PR 14-MAY-1999; 99US-0132646P.
PR 14-MAY-1999; 99US-0132647P.
PR 14-MAY-1999; 99US-0132648P.
PR 14-MAY-1999; 99US-0132649P.
PR 14-MAY-1999; 99US-0132650P.
PR 14-MAY-1999; 99US-0132651P.
PR 14-MAY-1999; 99US-0132652P.
PR 14-MAY-1999; 99US-0132653P.
PR 14-MAY-1999; 99US-0132654P.
PR 14-MAY-1999; 99US-0132655P.
PR 14-MAY-1999; 99US-0132656P.
PR 14-MAY-1999; 99US-0132657P.
PR 14-MAY-1999; 99US-0132658P.
PR 14-MAY-1999; 99US-0132659P.
PR 14-MAY-1999; 99US-0132660P.
PR 14-MAY-1999; 99US-0132661P.
PR 14-MAY-1999; 99US-0132662P.
PR 14-MAY-1999; 99US-0132663P.
PR 14-MAY-1999; 99US-0132664P.
PR 14-MAY-1999; 99US-0132665P.
PR 14-MAY-1999; 99US-0132666P.
PR 14-MAY-1999; 99US-0132667P.
PR 14-MAY-1999; 99US-0132668P.
PR 14-MAY-1999; 99US-0132669P.
PR 14-MAY-1999; 99US-0132670P.
PR 14-MAY-1999; 99US-0132671P.
PR 14-MAY-1999; 99US-0132672P.
PR 14-MAY-1999; 99US-0132673P.
PR 14-MAY-1999; 99US-0132674P.
PR 14-MAY-1999; 99US-0132675P.
PR 14-MAY-1999; 99US-0132676P.
PR 14-MAY-1999; 99US-0132677P.
PR 14-MAY-1999; 99US-0132678P.
PR 14-MAY-1999; 99US-0132679P.
PR 14-MAY-1999; 99US-0132680P.
PR 14-MAY-1999; 99US-0132681P.
PR 14-MAY-1999; 99US-0132682P.
PR 14-MAY-1999; 99US-0132683P.
PR 14-MAY-1999; 99US-0132684P.
PR 14-MAY-1999; 99US-0132685P.
PR 14-MAY-1999; 99US-0132686P.
PR 14-MAY-1999; 99US-0132687P.
PR 14-MAY-1999; 99US-0132688P.
PR 14-MAY-1999; 99US-0132689P.
PR 14-MAY-1999; 99US-0132690P.
PR 14-MAY-1999; 99US-0132691P.
PR 14-MAY-1999; 99US-0132692P.
PR 14-MAY-1999; 99US-0132693P.
PR 14-MAY-1999; 99US-0132694P.
PR 14-MAY-1999; 99US-0132695P.
PR 14-MAY-1999; 99US-0132696P.
PR 14-MAY-1999; 99US-0132697P.
PR 14-MAY-1999; 99US-0132698P.
PR 14-MAY-1999; 99US-0132699P.
PR 14-MAY-1999; 99US-0132700P.
PR 14-MAY-1999; 99US-0132701P.
PR 14-MAY-1999; 99US-0132702P.
PR 14-MAY-1999; 99US-0132703P.
PR 14-MAY-1999; 99US-0132704P.
PR 14-MAY-1999; 99US-0132705P.
PR 14-MAY-1999; 99US-0132706P.
PR 14-MAY-1999; 99US-0132707P.
PR 14-MAY-1999; 99US-0132708P.
PR 14-MAY-1999; 99US-0132709P.
PR 14-MAY-1999; 99US-0132710P.
PR 14-MAY-1999; 99US-0132711P.
PR 14-MAY-1999; 99US-0132712P.
PR 14-MAY-1999; 99US-0132713P.
PR 14-MAY-1999; 99US-0132714P.
PR 14-MAY-1999; 99US-0132715P.
PR 14-MAY-1999; 99US-0132716P.
PR 14-MAY-1999; 99US-0132717P.
PR 14-MAY-1999; 99US-0132718P.
PR 14-MAY-1999; 99US-0132719P.
PR 14-MAY-1999; 99US-0132720P.
PR 14-MAY-1999; 99US-0132721P.
PR 14-MAY-1999; 99US-0132722P.
PR 14-MAY-1999; 99US-0132723P.
PR 14-MAY-1999; 99US-0132724P.
PR 14-MAY-1999; 99US-0132725P.
PR 14-MAY-1999; 99US-0132726P.
PR 14-MAY-1999; 99US-0132727P.
PR 14-MAY-1999; 99US-0132728P.
PR 14-MAY-1999; 99US-0132729P.
PR 14-MAY-1999; 99US-0132730P.
PR 14-MAY-1999; 99US-0132731P.
PR 14-MAY-1999; 99US-0132732P.
PR 14-MAY-1999; 99US-0132733P.
PR 14-MAY-1999; 99US-0132734P.
PR 14-MAY-1999; 99US-0132735P.
PR 14-MAY-1999; 99US-0132736P.
PR 14-MAY-1999; 99US-0132737P.
PR 14-MAY-1999; 99US-0132738P.
PR 14-MAY-1999; 99US-0132739P.
PR 14-MAY-1999; 99US-0132740P.
PR 14-MAY-1999; 99US-0132741P.
PR 14-MAY-1999; 99US-0132742P.
PR 14-MAY-1999; 99US-0132743P.
PR 14-MAY-1999; 99US-0132744P.
PR 14-MAY-1999; 99US-0132745P.
PR 14-MAY-1999; 99US-0132746P.
PR 14-MAY-1999; 99US-0132747P.
PR 14-MAY-1999; 99US-0132748P.
PR 14-MAY-1999; 99US-0132749P.
PR 14-MAY-1999; 99US-0132750P.
PR 14-MAY-1999; 99US-0132751P.
PR 14-MAY-1999; 99US-0132752P.
PR 14-MAY-1999; 99US-0132753P.
PR 14-MAY-1999; 99US-0132754P.
PR 14-MAY-1999; 99US-0132755P.
PR 14-MAY-1999; 99US-0132756P.
PR 14-MAY-1999; 99US-0132757P.
PR 14-MAY-1999; 99US-0132758P.
PR 14-MAY-1999; 99US-0132759P.
PR 14-MAY-1999; 99US-0132760P.
PR 14-MAY-1999; 99US-0132761P.
PR 14-MAY-1999; 99US-0132762P.
PR 14-MAY-1999; 99US-0132763P.
PR 14-MAY-1999; 99US-0132764P.
PR 14-MAY-1999; 99US-0132765P.
PR 14-MAY-1999; 99US-0132766P.
PR 14-MAY-1999; 99US-0132767P.
PR 14-MAY-1999; 99US-0132768P.
PR 14-MAY-1999; 99US-0132769P.
PR 14-MAY-1999; 99US-0132770P.
PR 14-MAY-1999; 99US-0132771P.
PR 14-MAY-1999; 99US-0132772P.
PR 14-MAY-1999; 99US-0132773P.
PR 14-MAY-1999; 99US-0132774P.
PR 14-MAY-1999; 99US-0132775P.
PR 14-MAY-1999; 99US-0132776P.
PR 14-MAY-1999; 99US-0132777P.
PR 14-MAY-1999; 99US-0132778P.
PR 14-MAY-1999; 99US-0132779P.
PR 14-MAY-1999; 99US-0132780P.
PR 14-MAY-1999; 99US-0132781P.
PR 14-MAY-1999; 99US-0132782P.
PR 14-MAY-1999; 99US-0132783P.
PR 14-MAY-1999; 99US-0132784P.
PR 14-MAY-1999; 99US-0132785P.
PR 14-MAY-1999; 99US-0132786P.
PR 14-MAY-1999; 99US-0132787P.
PR 14-MAY-1999; 99US-0132788P.
PR 14-MAY-1999; 99US-0132789P.
PR 14-MAY-1999; 99US-0132790P.
PR 14-MAY-1999; 99US-0132791P.
PR 14-MAY-1999; 99US-0132792P.
PR 14-MAY-1999; 99US-0132793P.
PR 14-MAY-1999; 99US-0132794P.
PR 14-MAY-1999; 99US-0132795P.
PR 14-MAY-1999; 99US-0132796P.
PR 14-MAY-1999; 99US-0132797P.
PR 14-MAY-1999; 99US-0132798P.
PR 14-MAY-1999; 99US-0132799P.
PR 14-MAY-1999; 99US-0132800P.
PR 14-MAY-1999; 99US-0132801P.
PR 14-MAY-1999; 99US-0132802P.
PR 14-MAY-1999; 99US-0132803P.
PR 14-MAY-1999; 99US-0132804P.
PR 14-MAY-1999; 99US-0132805P.
PR 14-MAY-1999; 99US-0132806P.
PR 14-MAY-1999; 99US-0132807P.
PR 14-MAY-1999; 99US-0132808P.
PR 14-MAY-1999; 99US-0132809P.
PR 14-MAY-1999; 99US-0132810P.
PR 14-MAY-1999; 99US-0132811P.
PR 14-MAY-1999; 99US-0132812P.
PR 14-MAY-1999; 99US-0132813P.
PR 14-MAY-1999; 99US-0132814P.
PR 14-MAY-1999; 99US-0132815P.
PR 14-MAY-1999; 99US-0132816P.
PR 14-MAY-1999; 99US-0132817P.
PR 14-MAY-1999; 99US-0132818P.
PR 14-MAY-1999; 99US-0132819P.
PR 14-MAY-1999; 99US-0132820P.
PR 14-MAY-1999; 99US-0132821P.
PR 14-MAY-1999; 99US-0132822P.
PR 14-MAY-1999; 99US-0132823P.
PR 14-MAY-1999; 99US-0132824P.
PR 14-MAY-1999; 99US-0132825P.
PR 14-MAY-1999; 99US-0132826P.
PR 14-MAY-1999; 99US-0132827P.
PR 14-MAY-1999; 99US-0132828P.
PR 14-MAY-1999; 99US-0132829P.
PR 14-MAY-1999; 99US-0132830P.
PR 14-MAY-1999; 99US-0132831P.
PR 14-MAY-1999; 99US-0132832P.
PR 14-MAY-1999; 99US-0132833P.
PR 14-MAY-1999; 99US-0132834P.
PR 14-MAY-1999; 99US-0132835P.
PR 14-MAY-1999; 99US-0132836P.
PR 14-MAY-1999; 99US-0132837P.
PR 14-MAY-1999; 99US-0132838P.
PR 14-MAY-1999; 99US-0132839P.
PR 14-MAY-1999; 99US-0132840P.
PR 14-MAY-1999; 99US-0132841P.
PR 14-MAY-1999; 99US-0132842P.
PR 14-MAY-1999; 99US-0132843P.
PR 14-MAY-1999; 99US-0132844P.
PR 14-MAY-1999; 99US-0132845P.
PR 14-MAY-1999; 99US-0132846P.
PR 14-MAY-1999; 99US-0132847P.
PR 14-MAY-1999; 99US-0132848P.
PR 14-MAY-1999; 99US-0132849P.
PR 14-MAY-1999; 99US-0132850P.
PR 14-MAY-1999; 99US-0132851P.
PR 14-MAY-1999; 99US-0132852P.
PR 14-MAY-1999; 99US-0132853P.
PR 14-MAY-1999; 99US-0132854P.
PR 14-MAY-1999; 99US-0132855P.
PR 14-MAY-1999; 99US-0132856P.
PR 14-MAY-1999; 99US-0132857P.
PR 14-MAY-1999; 99US-0132858P.
PR 14-MAY-1999; 99US-0132859P.
PR 14-MAY-1999; 99US-0132860P.
PR 14-MAY-1999; 99US-0132861P.
PR 14-MAY-1999; 99US-0132862P.
PR 14-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0132864P.
PR 14-MAY-1999; 99US-0132865P.
PR 14-MAY-1999; 99US-0132866P.
PR 14-MAY-1999; 99US-0132867P.
PR 14-MAY-1999; 99US-0132868P.
PR 14-MAY-1999; 99US-0132869P.
PR 14-MAY-1999; 99US-0132870P.
PR 14-MAY-1999; 99US-0132871P.
PR 14-MAY-1999; 99US-0132872P.
PR 14-MAY-1999; 99US-0132873P.
PR 14-MAY-1999; 99US-0132874P.
PR 14-MAY-1999; 99US-0132875P.
PR 14-MAY-1999; 99US-0132876P.
PR 14-MAY-1999; 99US-0132877P.
PR 14-MAY-1999; 99US-0132878P.
PR 14-MAY-1999; 99US-0132879P.
PR 14-MAY-1999; 99US-0132880P.
PR 14-MAY-1999; 99US-0132881P.
PR 14-MAY-1999; 99US-0132882P.
PR 14-MAY-1999; 99US-0132883P.
PR 14-MAY-1999; 99US-0132884P.
PR 14-MAY-1999; 99US-0132885P.
PR 14-MAY-1999; 99US-0132886P.
PR 14-MAY-1999; 99US-0132887P.
PR 14-MAY-1999; 99US-0132888P.
PR 14-MAY-1999; 99US-0132889P.
PR 14-MAY-1999; 99US-0132890P.
PR 14-MAY-1999; 99US-0132891P.
PR 14-MAY-1999; 99US-0132892P.
PR 14-MAY-1999; 99US-0132893P.
PR 14-MAY-1999; 99US-0132894P.
PR 14-MAY-1999; 99US-0132895P.
PR 14-MAY-1999; 99US
```



PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-015065P.  
 PR 27-AUG-1999; 99US-015066P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154779P.  
 PR 20-SEP-1999; 99US-0155133P.  
 PR 22-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156459P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159332P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 18-OCT-1999; 99US-0159638P.  
 PR 21-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.

PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 92.5%; Score 49; DB 3; Length 81;  
 Best Local Similarity 80.0%; Pred. No. 0.96;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10  
 |||:|||||:  
 Db 6 RLQDPPAGI 15

RESULT 7  
 AAG19026  
 ID AAG19026 standard; protein; 90 AA.  
 XX AAG19026;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Zea mays protein fragment SEQ ID NO: 20668.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence; corn.  
 XX  
 OS Zea mays subsp. mays.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144083P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 20-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 06-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149910P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.

PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.5%; Score 49; DB 3; Length 90;  
Best Local Similarity 80.0%; Pred. No. 1.1;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10  
DB 15 RLQDPPAGI 24  
|||:||||:

## RESULT 8

ID ABM74252 standard; protein; 141 AA.

XX AC ABM74252;  
XX DT 17-OCT-2003 (first entry)

DE DNA clone originating in barley containing SNP sequence #662.

XX KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.

OS Hordeum vulgare.

PN WO2003057877-A1.

XX PD 17-JUL-2003.

XX PF 16-DEC-2002; 2002WO-IB005403.

XX PR 20-DEC-2001; 2001JP-00387059.

XX PR 20-DEC-2001; 2001JP-00387131.

XX PR 20-DEC-2001; 2001JP-00403299.

XX PR 20-DEC-2001; 2001JP-00403300.

XX PR 27-SEP-2002; 2002JP-00327515.

XX PA (UYN1-) UNIV JAPAN OKAYAMA.

XX PI Sato K, Takeda K, Kohara Y;

XX DR WPI; 2003-587127/55.

XX PT Single nucleotide polymorphism sites in barley varieties and DNA  
PT sequences containing them for analysis and identification of barley  
PT varieties and production of barley transformants with desired  
PT characteristics.  
XX Disclosure; SEQ ID XX; 284pp; Japanese.

XX CC The present invention relates to oligonucleotide clones originating in  
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms  
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
CC varieties, identification of particular varieties and genotype-phenotype  
CC analysis, isolation of specific genes and creation of new varieties by  
CC transformation of barley varieties with them and production of new barley  
CC varieties with desired properties. The present sequence represents an  
CC oligonucleotide clone sequence featured in the specification. The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published-pct-sequences

XX SQ Sequence 141 AA;

Query Match 92.5%; Score 49; DB 7; Length 141;  
Best Local Similarity 80.0%; Pred. No. 1.6;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10  
DB 4 RLQDPPAGI 13  
|||:||||:

## RESULT 9

AAG10241

ID AAG10241 standard; protein; 143 AA.

XX AC AAG10241;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 8479.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

XX PR 23-APR-1999; 99US-0130891P.

XX PR 28-APR-1999; 99US-0131449P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 30-APR-1999; 99US-0132407P.

XX PR 04-MAY-1999; 99US-0132484P.

XX PR 05-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 06-MAY-1999; 99US-0132487P.

XX PR 07-MAY-1999; 99US-0132863P.

XX PR 11-MAY-1999; 99US-0134256P.

XX PR 14-MAY-1999; 99US-0134218P.

XX PR 14-MAY-1999; 99US-0134219P.

XX PR 14-MAY-1999; 99US-0134221P.

XX PR 14-MAY-1999; 99US-0134370P.

XX PR 18-MAY-1999; 99US-0134768P.

XX PR 19-MAY-1999; 99US-0134941P.

XX PR 20-MAY-1999; 99US-0135124P.

XX PR 21-MAY-1999; 99US-0135353P.

XX PR 24-MAY-1999; 99US-0135629P.

XX PR 25-MAY-1999; 99US-0136021P.

XX PR 27-MAY-1999; 99US-0136392P.

XX PR 28-MAY-1999; 99US-0136782P.

XX PR 01-JUN-1999; 99US-0137222P.

XX PR 03-JUN-1999; 99US-0137528P.

XX PR 04-JUN-1999; 99US-0137502P.

XX PR 07-JUN-1999; 99US-0137724P.

XX PR 08-JUN-1999; 99US-0138094P.

XX PR 10-JUN-1999; 99US-0138540P.

XX PR 10-JUN-1999; 99US-0138847P.

XX PR 14-JUN-1999; 99US-0139119P.

XX PR 16-JUN-1999; 99US-0139452P.

XX PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154775P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.5%; Score 49; DB 3; Length 143;  
Best Local Similarity 80.0%; Pred. No. 1.6;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLOEDPPAGV 10  
| | | | | | | | | |  
Db 6 RLQODPPAGI 15

RESULT 10



```

PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 06-OCT-1999; 99US-0157753P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159332P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 92.5%; Score 49; DB 3; Length 143;
Best Local Similarity 80.0%; Pred. No. 1.6;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQDPPAGV 10
DB 6 RLQDPPAGI 15

RESULT 11
AAG32531
ID AAG32531 standard; protein; 143 AA.
XX
AC AAG32531;
XX
DT 17-OCT-2000 (first entry)
XX
```

```

DE Arabidopsis thaliana protein fragment SEQ ID NO: 39262.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN EF1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129645P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
```

```
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145227P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147491P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
```

```
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 92.5%; Score 49; DB 3; Length 143;  
Best Local Similarity 80.0%; Pred. No. 1.6;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
|||:|||||:  
Db 6 RLQODPPAGI 15

## RESULT 12

ABBS8767  
ID ABBS8767 standard; protein; 151 AA.

XX AC ABBS8767;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 3093.

XX KW Drosophila; developmental biology; cell signalling; insecticide;  
pharmaceutical.

XX OS Drosophila melanogaster.

PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US009231.  
 XX  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX WPI; 2001-656860/75.  
 DR N-PSDB; ABL02870.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 XX Disclosure; SEQ ID NO 3093; 21pp + Sequence Listing; English.  
 PS  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 151 AA;  
 SQ  
 Query Match 92.5%; Score 49; DB 4; Length 151;  
 Best Local Similarity 90.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 RLQEDPPAGV 10  
 Db |||||  
 15 RLQEDPPTGV 24  
 RESULT 13  
 ID AAR67277 standard; protein; 152 AA.  
 XX  
 AC AAR67277;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 11-AUG-1995 (first entry)  
 XX  
 DE Ubiquitin-conjugating enzyme UBC1.  
 XX  
 XX Ubiquitin-conjugating enzyme; UBC; E2; ubiquitin carrier protein;  
 KW protein degradation; proteolysis.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX EP626450-A2.  
 PN  
 XX 30-NOV-1994.  
 PD  
 XX 31-MAY-1994; 94EP-00303903.  
 XX  
 XX 28-MAY-1993; 93US-00070157.  
 XX  
 XX (WISC ) WISCONSIN ALUMNI RES FOUND.  
 PA  
 XX Vierstra RD, Gosink MM;  
 PI  
 XX WPI; 1995-000934/01.  
 DR

DR N-PSDB; AAQ75293.  
 XX  
 PT New DNA encoding new fusion protein of E2 protein and target protein -  
 PT able to transfer ubiquitin specifically to the target to induce its  
 PT degradation in a host, e.g. for inhibiting viral infection in plants or  
 PT mammals.  
 XX  
 PS Claim 17; Page 12-13; 30pp; English.  
 XX  
 CC UBC1 is an E2 isolated from Arabidopsis thaliana, while UBC4 is an E2  
 CC isolated from wheat. Native UBC4 E2 is capable, in vitro, of binding to  
 CC the target protein alone, without the need for an ubiquitin protein  
 CC ligase, or E3, molecule. UBC1 requires the presence of the ubiquitin  
 CC protein ligase, or E3 molecule, in order to direct the ubiquitin addition  
 CC to a target molecule. UBC4 possesses a tail region at the carboxyl  
 CC terminus, while UBC1 possess no such tail. It is postulated in the patent  
 CC that the transfer of a tail alone will not transform UBC1 into an E3-  
 CC independent E2 isoform; there must also be a spacer region between the  
 CC carboxyl terminus of the tail-less E2 and the tail. (Updated on 25-MAR-  
 CC 2003 to correct PN field.)  
 XX  
 SQ Sequence 152 AA;  
 Query Match 92.5%; Score 49; DB 2; Length 152;  
 Best Local Similarity 80.0%; Pred. No. 1.7;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 RLQEDPPAGV 10  
 Db |||||  
 15 RLQODPPAGI 24  
 RESULT 14  
 ID AAR79654 standard; protein; 152 AA.  
 XX  
 AC AAR79654;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 06-DEC-1995 (first entry)  
 XX  
 DE Human E2 protein.  
 XX  
 KW Ubiquitin-conjugating enzyme; E2 protein; cell cycle; cell proliferation;  
 KW cancer; psoriasis; fibrosis.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9518974-A2.  
 PN  
 XX 13-JUL-1995.  
 PD  
 XX 04-JAN-1995; 95WO-US000164.  
 XX  
 XX 04-JAN-1994; 94US-00176937.  
 PR 23-MAY-1994; 94US-00247904.  
 PR 27-MAY-1994; 94US-00250795.  
 PR 13-SEP-1994; 94US-00305520.  
 XX  
 XX (MITO-) MITOTIX INC.  
 PA  
 XX Draetta G, Rolfe M, Eckstein JW, Cottarel G, Gyuris J;  
 PI  
 XX WPI; 1995-255137/33.  
 DR N-PSDB; AAQ97844.  
 XX  
 XX Identifying inhibitors of ubiquitin mediated proteolysis of cell cycle  
 PT regulatory proteins - also new ubiquitin conjugating enzymes, their  
 PT related nucleic acid, vectors, antibodies etc., useful for regulating  
 PT e.g. cell proliferation.  
 XX  
 PS Disclosure; Page 99-100; 157pp; English.  
 XX



CC Human E2 cDNA (given in AAQ97844) was amplified from a HeLa cell cDNA  
 CC library using the primers given in AAQ97842-43. The gene was subcloned  
 CC into a baculovirus or pEX vector for expression of recombinant E2 in Sf9  
 CC insect or E. coli cells for use as a component of an in vitro ubiquitin  
 CC conjugating system. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 152 AA;

Query Match 92.5%; Score 49; DB 2; Length 152;  
 Best Local Similarity 90.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10  
 |||||  
 Db 15 RLQEDPPVGV 24

## RESULT 15

AAV39967  
 ID AAV39967 standard; protein; 152 AA.

XX  
 AC AAV39967;

XX  
 DT 15-DEC-1999 (first entry)

XX  
 DE Human rad6 homolog protein sequence.

XX Ubiquitin conjugating enzyme; UbCE; ubiquitin-mediated proteolysis;  
 KW cell-cycle regulatory protein; ubiquitination inhibitor; atherosclerosis;  
 KW proliferative disorder; cancer; restenosis; tissue connective disorder;  
 KW wound healing; fibrosis disorder; rheumatoid arthritis; scleroderma;  
 KW insulin dependent diabetes mellitus; glomerulonephritis; cirrhosis;  
 KW diagnosis; therapy; rad6; ss.

XX  
 OS Homo sapiens.

XX  
 PN US5968761-A.

XX  
 PD 19-OCT-1999.

XX  
 PF 07-JUN-1995; 95US-00486663.

XX  
 PR 04-JAN-1994; 94US-00176937.

XX  
 PR 23-MAY-1994; 94US-00247904.

XX  
 PR 13-SEP-1994; 94US-00250795.

XX  
 PA (MITO-) MITOTIX INC.

XX  
 PI Chiu MI, Cottarel G, Berlin V, Damagnez V, Draetta G, Rolfe M;

XX  
 DR WPI; 1999-590402/50.

XX  
 DR N-PSDB; AAZ27567.

XX  
 PT Identifying ubiquitination inhibitors using novel ubiquitin conjugating  
 PT enzymes.

XX  
 PS Example 2; Col 89-90; 61pp; English.

XX This sequence is the human rad6 protein. The invention relates to assays  
 CC for identifying an inhibitor of ubiquitin-mediated proteolysis of a cell-  
 CC cycle regulatory protein comprising contacting a candidate agent with an  
 CC ubiquitin-conjugating system and measuring the level of ubiquitination.  
 CC The ubiquitin-conjugating system comprises: (a) a reconstituted protein  
 CC mixture including a ubiquitin conjugating enzyme (UbCE) produced by the  
 CC expression of a nucleic acid which hybridizes under high stringency  
 CC conditions to human UbCE, Candida albicans UbCE, or Schizosaccharomyces  
 CC pombe UbCE coding sequences; (b) a regulatory protein; and (c) ubiquitin.  
 CC The polynucleotides are useful for identifying ubiquitination inhibitors.  
 CC The polynucleotides, polypeptides, antisense compounds and antibodies  
 CC against them may also be useful for the treatment and/or diagnosis of  
 CC proliferative disorders (e.g. cancer, atherosclerosis, or restenosis),  
 CC tissue connective disorders, controlling wound healing, and disorders

CC characterized by fibrosis (e.g. rheumatoid arthritis, insulin dependent  
 CC diabetes mellitus, glomerulonephritis, cirrhosis, and scleroderma)  
 XX  
 SQ Sequence 152 AA;

Query Match 92.5%; Score 49; DB 2; Length 152;  
 Best Local Similarity 90.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10  
 |||||  
 Db 15 RLQEDPPVGV 24

Search completed: November 18, 2005, 01:00:54  
 Job time : 59.1936 secs

**This Page Blank (usp10)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:11:06 ; Search time 9.89247 Seconds  
(without alignments)  
97.263 Million cell updates/sec

Title: US-10-006-177-2

Perfect score: 53

Sequence: 1 RLQEDPPAGV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79:\*

2: PIR1:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	100.0	152	2 A41222	ubiquitin-protein
2	49	92.5	151	2 A39392	RAD6 DNA-repair ho
3	49	92.5	152	2 B41222	ubiquitin-protein
4	49	92.5	152	2 A42416	ubiquitin carrier
5	49	92.5	152	2 I51913	14 kDa ubiquitin c
6	49	92.5	152	2 S43783	ubiquitin-conjugat
7	49	92.5	152	2 S43781	ubiquitin-conjugat
8	48	90.6	151	2 S12529	ubiquitin-conjugat
9	48	90.6	151	2 T45220	ubiquitin-protein
10	45	84.9	151	2 S71430	DNA repair protein
11	45	84.9	151	2 T51931	hypothetical prote
12	44	83.0	150	2 S43782	ubiquitin-conjugat
13	41	77.4	760	2 JQ1383	hypothetical 84k p
14	39	73.6	205	2 T32959	hypothetical prote
15	39	73.6	635	2 D64835	ABC-type transport
16	39	73.6	635	2 G85621	hypothetical prote
17	39	73.6	635	2 A90758	hypothetical prote
18	39	73.6	637	2 A80173	ABC transporter AT
19	38	71.7	635	2 A10625	ABC transporter AT
20	38	71.7	1439	2 T27110	hypothetical prote
21	37	69.8	127	2 A45944	major sperm protei
22	37	69.8	443	2 B75165	signal recognition
23	37	69.8	445	2 H71176	probable signal re
24	37	69.8	1597	2 S55144	RUB1 protein - yea
25	36	67.9	345	2 D72536	hypothetical prote
26	36	67.9	540	2 T00646	hypothetical prote
27	36	67.9	1238	2 T15919	hypothetical prote
28	35	66.0	77	2 F88165	protein ZK1248.4 [
29	35	66.0	95	2 T34500	hypothetical prote

protein C04G2.4 [i  
hypothetical prote  
hypothetical prote  
protein C34F11.6 [i  
protein ZK546.6 [i  
protein K05F1.7 [i  
protein C34F11.4 [i  
protein msp-19 [im  
protein ZK1248.6 [i  
protein msp-65 [im  
protein MSP-31 [im  
protein F58A6.8 [i  
protein msp-142 [i  
protein msp-59 [im  
protein msp-51 [im  
protein msp-113 [i

#### ALIGNMENTS

##### RESULT 1

A41222

ubiquitin-protein ligase (EC 6.3.2.19) E2A - human

N:Alternate names: ubiquitin-conjugating enzyme HHR6A

C:Species: Homo sapiens (man)

C>Date: 19-Jun-1992 #sequence\_revision 19-Jun-1992 #text\_change 09-Jul-2004

C:Accession: A41222

R:Koken, M.H.M.; Reynolds, P.; Jaspers-Dekker, I.; Prakash, L.; Prakash, S.; Bootsma, D.

Proc. Natl. Acad. Sci. U.S.A. 88, 8865-8869, 1991

A:Title: Structural and functional conservation of two human homologs of the yeast DNA R

A:Reference number: A41222; MUID:92020951; PMID:1717990

A:Accession: A41222

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-152 <KOK>

A:Cross-references: UNIPROT:P49459; GB:M74524; NID:gl84043; PIDN:AAA35981.1; PID:gl84044

C:Genetics:

A:Gene: GDB:UBB2A; UBC2; HHR6A

A:Map position: Xq24-Xq25

C:Superfamily: human ubiquitin-protein ligase E2

C:Keywords: ligase; protein degradation; ubiquitination

F:88/Active site: Cys #status predicted

Query Match 100.0%; Score 53; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
|||  
Db 15 RLQEDPPAGV 24

##### RESULT 2

A39392

RAD6 DNA-repair homolog Dhr6 - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 28-Feb-1992 #sequence\_revision 28-Feb-1992 #text\_change 09-Jul-2004

C:Accession: A39392

R:Koken, M.; Reynolds, P.; Bootsma, D.; Hoeijmakers, J.; Prakash, S.; Prakash, L.

Proc. Natl. Acad. Sci. U.S.A. 88, 3832-3836, 1991

A:Title: Dhr6, a Drosophila homolog of the yeast DNA-repair gene RAD6.

A:Reference number: A39392; MUID:91219466; PMID:1902572

A:Accession: A39392

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-151 <KOK>

A:Cross-references: UNIPROT:P25153; GB:M63791; NID:gl56731; PID:gl56732

C:Genetics:

A:Gene: FlyBase:UbcD6

A:Cross-references: FlyBase:FBgn0004436

C:Superfamily: human ubiquitin-protein ligase E2

Query Match 92.5%; Score 49; DB 2; Length 151;  
Best Local Similarity 90.0%; Pred. No. 0.061; Length 151;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPGV 24  
|||||

RESULT 3  
B41222  
ubiquitin-protein ligase (EC 6.3.2.19) E2B - human  
N;Alternate names: ubiquitin carrier protein E2; ubiquitin-conjugating enzyme HHR6B  
C;Species: Homo sapiens (man)  
C;Date: 19-Jun-1992 #sequence\_revision 19-Jun-1992 #text\_change 09-Jul-2004  
C;Accession: B41222; S12530; S17516  
R;Koken, M.H.M.; Reynolds, P.; Jaspers-Dekker, I.; Prakash, L.; Prakash, S.; Bootsma, D.  
Proc. Natl. Acad. Sci. U.S.A. 88, 8865-8869, 1991  
A;Title: Structural and functional conservation of two human homologs of the yeast DNA x  
A;Reference number: A41222; MUID:92020951; PMID:1717990  
A;Accession: B41222  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-152 <KOK>  
A;Cross-references: UNIPROT:P23567; GB:M74525; NID:g184045; PIDN:AAA35982.1; PID:g184046  
R;Schneider, R.; Eckerskorn, C.; Lottspeich, F.; Schweiger, M.  
EMBO J. 9, 1431-1435, 1990  
A;Title: The human ubiquitin carrier protein E2(M(r) = 17000) is homologous to the yeast  
A;Reference number: S12530; MUID:90228340; PMID:2158443  
A;Accession: S12530  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-152 <SCH>  
A;Cross-references: GB:X53251; NID:g30953; PIDN:CAA37339.1; PID:g30954  
R;Woffendin, C.; Chen, Z.; Staskus, K.; Retzel, E.F.; Plagemann, P.G.W.  
Biochim. Biophys. Acta 1090, 81-85, 1991  
A;Title: Mammalian mRNAs encoding protein closely related to ubiquitin-conjugating enzym  
A;Reference number: S17516; MUID:91355233; PMID:1893845  
A;Accession: S17516  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-21,'C',24-40,'I',42-53,'R',55-152 <WOF>  
C;Genetics:  
A;Gene: GDB:UBB2B; USC2; HHR6B  
A;Cross-references: GDB:I28627; OMIM:179095  
A;Map position: 5q23-5q31  
C;Superfamily: human ubiquitin-protein ligase E2  
C;Keywords: ligase; protein degradation; ubiquitination  
F;88/Active site: Cys #status predicted

Query Match 92.5%; Score 49; DB 2; Length 152;  
Best Local Similarity 90.0%; Pred. No. 0.061; Length 152;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPGV 24  
|||||

RESULT 4  
A42416  
ubiquitin carrier protein E2 - rabbit  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C;Accession: A42416  
R;Wing, S.S.; Dumas, F.; Banville, D.  
J. Biol. Chem. 267, 6495-6501, 1992  
A;Title: A rabbit reticulocyte ubiquitin carrier protein that supports ubiquitin-depende  
A;Reference number: A42416; MUID:92202189; PMID:1313008  
A;Accession: A42416  
A;Status: preliminary  
A;Molecule type: mRNA

A;Residues: 1-152 <WIN>  
A;Cross-references: UNIPROT:P23567; GB:M62387; NID:g165779; PIDN:AAA31492.1; PID:g165780  
A;Experimental source: reticulocyte  
A;Note: sequence extracted from NCBI backbone (NCBIN:89551, NCBI:P:89553)  
C;Superfamily: human ubiquitin-protein ligase E2

Query Match 92.5%; Score 49; DB 2; Length 152;  
Best Local Similarity 90.0%; Pred. No. 0.061; Length 152;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPGV 24  
|||||

RESULT 5  
I51913  
14 kDa ubiquitin conjugating enzyme - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I51913  
R;Wing, S.S.; Banville, D.  
Am. J. Physiol. 267, E39-E48, 1994  
A;Title: The 14 kDa ubiquitin conjugating enzyme: structure of the rat gene and regulati  
A;Reference number: I51913; MUID:94324482; PMID:8048511  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-152 <RES>  
A;Cross-references: UNIPROT:P23567; EMBL:U04308; NID:g476115; PIDN:AAB60669.1; PID:g4761  
C;Genetics:  
A;Introns: 15/2; 42/2; 51/1; 81/1; 110/3  
C;Superfamily: human ubiquitin-protein ligase E2

Query Match 92.5%; Score 49; DB 2; Length 152;  
Best Local Similarity 90.0%; Pred. No. 0.061; Length 152;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPGV 24  
|||||

RESULT 6  
S43783  
ubiquitin-conjugating enzyme USC2 - Arabidopsis thaliana  
N;Alternate names: protein T20F6.10; ubiquitin-conjugating enzyme E2  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S43783; T00852; E84440  
R;Sullivan, M.L.; Carpenter, T.B.; Vierstra, R.D.  
Plant Mol. Biol. 24, 651-661, 1994  
A;Title: Homologues of wheat ubiquitin-conjugating enzymes - TaUBC1 and TaUBC4 are encod  
A;Reference number: S43781; MUID:94207190; PMID:8155884  
A;Accession: S43783  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-152 <SUL>  
A;Cross-references: UNIPROT:P42745; EMBL:L19353; NID:g431263; PIDN:AAA32899.1; PID:g4312  
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, March 1998  
A;Description: Arabidopsis thaliana chromosome II BAC T20F6 genomic sequence.  
A;Reference number: Z14206  
A;Accession: T00852  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-152 <ROU>  
A;Cross-references: EMBL:AC002521; NID:g2947056; PID:g2947065  
A;Experimental source: cultivar Columbia  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E84440

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-152 <STO>

A:Cross-references: GB:AE002093; NID:g2947065; PIDN:AAC05346.1; GSPDB:GN00139

C:Genetics;

A:Gene: UBC2; At2g02760; T20R6.10

A:Map position: 2

A:Introns: 42/2; 51/1; 81/1; 110/3

C:Superfamily: human ubiquitin-protein ligase E2

Query Match 92.5%; Score 49; DB 2; Length 152;

Best Local Similarity 80.0%; Pred. No. 0.061;

Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10

|||:|||||:

Db 15 RLQDPPAGI 24

RESULT 7

S43781

ubiquitin-conjugating enzyme UBC1 - Arabidopsis thaliana

A:Alternate names: ubiquitin-conjugating enzyme E2

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 09-Jul-2004

C:Accession: S43781

R:Sullivan, M.L.; Carpenter, T.B.; Vierstra, R.D.

Plant Mol. Biol. 24, 651-661, 1994

A:Title: Homologues of wheat ubiquitin-conjugating enzymes - TaUBC1 and TaUBC4 are encoded

A:Reference number: S43781; MUID:94207190; PMID:8155884

A:Accession: S43781

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-152 <SUL>

A:Cross-references: UNIPROT:P25865; EMBL:L19351; NID:G431259; PIDN:AAA32897.1; PID:G4312

C:Genetics:

A:Gene: UBC1

A:Introns: 42/2; 51/1; 81/1; 110/3

C:Superfamily: human ubiquitin-protein ligase E2

Query Match 92.5%; Score 49; DB 2; Length 152;

Best Local Similarity 80.0%; Pred. No. 0.061;

Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10

|||:|||||:

Db 15 RLQDPPAGI 24

RESULT 8

S12529

ubiquitin-conjugating enzyme rhp6 - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: protein SPAC18B11.07c

C:Species: Schizosaccharomyces pombe

C:Date: 21-Nov-1993 #sequence\_revision 12-Apr-1996 #text\_change 09-Jul-2004

C:Accession: S12529; T37907; S59845

R:Reynolds, P.; Koken, M.H.M.; Hoeijmakers, J.H.J.; Prakash, S.; Prakash, L.

EMBO J. 9, 1423-1430, 1990

A:Title: The rhp6(+) gene of Schizosaccharomyces pombe: a structural and functional homo

A:Reference number: S12529; MUID:90228339; PMID:2184030

A:Accession: S12529

A:Molecule type: DNA

A:Residues: 1-151 <REV>

A:Cross-references: UNIPROT:P23566; EMBL:Z50728; NID:g929886; PIDN:CAA90592.1; PID:g9298

R:Devlin, K.; Churcher, C.M.; Barrill, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, August 1995

A:Reference number: Z21753

A:Accession: T37907

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-151 <REV>

A:Cross-references: UNIPROT:P23566; EMBL:Z50728; NID:g929886; PIDN:CAA90592.1; PID:g9298

R:Devlin, K.; Churcher, C.M.; Barrill, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, August 1995

A:Reference number: Z21753

A:Accession: T37907

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-151 <DEV>

A:Cross-references: EMBL:Z50728; PIDN:CAA90592.1; PID:g929886; GSPDB:GN00066; SPDB:SPAC1

A:Experimental source: strain 972h-; cosmid c18B11

C:Genetics:

A:Gene: rhp6

A:Map position: 1L

A:Introns: 14/3; 36/2; 76/3; 111/2

C:Superfamily: human ubiquitin-protein ligase E2

C:Keywords: nucleus

Query Match 90.6%; Score 48; DB 2; Length 151;

Best Local Similarity 80.0%; Pred. No. 0.094;

Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10

|||:|||||:

Db 15 RMQDPPAGV 24

RESULT 9

T45220

ubiquitin-protein ligase (EC 6.3.2.19) rhp6 [imported] - fission yeast (Schizosaccharomy

C:Species: Schizosaccharomyces pombe

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T45220

R:Reynolds, P.; Koken, M.H.M.; Hoeijmakers, J.H.J.; Prakash, S.; Prakash, L.

EMBO J. 9, 1423-1430, 1990

A:Title: The rhp6(+) gene of Schizosaccharomyces pombe: a structural and functional homo

A:Reference number: S12529; MUID:90228339; PMID:2184030

A:Accession: T45220

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-151 <REV>

A:Cross-references: UNIPROT:P23566; EMBL:X53252; NID:g5045; PIDN:CAA37340.1; PID:g5046

C:Genetics:

A:Gene: rhp6

A:Introns: 14/3; 36/3; 76/3; 111/3

C:Superfamily: human ubiquitin-protein ligase E2

C:Keywords: ligase

Query Match 90.6%; Score 48; DB 2; Length 151;

Best Local Similarity 80.0%; Pred. No. 0.094;

Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10

|||:|||||:

Db 15 RMQDPPAGV 24

RESULT 10

S71430

DNA repair protein mus-8 - Neurospora crassa

C:Species: Neurospora crassa

C:Date: 12-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004

C:Accession: S71430

R:Soshi, T.; Sakuraba, Y.; Kaefler, E.; Inoue, H.

Curr. Genet. 30, 224-231, 1996

A:Title: The mus-8 gene of Neurospora crassa encodes a structural and functional homolog

A:Reference number: S71430; MUID:96337904; PMID:8753651

A:Accession: S71430

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-151 <SOS>

A:Cross-references: UNIPROT:P52493; EMBL:D78372; NID:gl100728; PIDN:BAA11380.1; PID:gl110

A:Note: the authors translated the codon GCC for residue 26 as Ser

C:Genetics:

A:Gene: mus-8

A:Introns: 14/3; 36/2; 111/2

C:Superfamily: human ubiquitin-protein ligase E2

Query Match 84.9%; Score 45; DB 2; Length 151;

Best Local Similarity 80.0%; Pred. No. 0.35;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 RLQEDPPAGV 10
      .|.|.|.|.|
Db      15 RMQTDPPAGV 24

RESULT 11
T51931
hypothetical protein Nhrad6 [imported] - Haematonectria haematococca
C/Species: Haematonectria haematococca
C/Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C/Accession: T51931
R/Wu, Q.; Aist, J.R.; Wirsal, S.G.; Turgeon, B.G.; Yoder, O.C.; Sandrock, T.
submitted to the EMBL Data Library, January 1997
A/Description: Nectria haematococca mating population VI Nhrad6 and Nhk1n1 genes.
A/Reference number: Z25871
A/Accession: T51931
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-151 <WUQ>
A/Cross-references: UNIPROT:P78717; EMBL:U86521; PIDN:AAB47850.1
A/Experimental source: strain T213 mating population VI
C/Genetics:
A/Note: Nhrad6
C/Superfamily: human ubiquitin-protein ligase E2

Query Match      84.9%; Score 45; DB 2; Length 151;
Best Local Similarity 80.0%; Pred. No. 0.35;
Matches      8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 RLQEDPPAGV 10
      .|.|.|.|.|
Db      15 RMQTDPPAGV 24

RESULT 12
S43782
ubiquitin-conjugating enzyme UBC3 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C/Accession: S43782
R/Sullivan, M.L.; Carpenter, T.B.; Vierstra, R.D.
Plant Mol. Biol. 24, 651-661, 1994
A/Title: Homologues of wheat ubiquitin-conjugating enzymes - TaUBC1 and TaUBC4 are encoded
A/Reference number: S43781; MUID:94207190; PMID:8155884
A/Accession: S43782
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-150 <SUL>
A/Cross-references: UNIPROT:P42746; EMBL:L19352; NID:G431261; PIDN:AAA32898.1; PID:G4312
C/Genetics:
A/Gene: UBC3
A/Introns: 42/2; 51/1; 81/1; 110/3
C/Superfamily: human ubiquitin-protein ligase E2

Query Match      83.0%; Score 44; DB 2; Length 150;
Best Local Similarity 70.0%; Pred. No. 0.53;
Matches      7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 RLQEDPPAGV 10
      .|.|.|.|.|
Db      15 RLQKDPPVGI 24

RESULT 13
JQ1383
hypothetical 84K protein - pea enation mosaic virus
C/Species: pea enation mosaic virus, PEMV
C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C/Accession: JQ1383
R/Demler, S.A.; de Zoeten, G.A.
J. Gen. Virol. 72, 1819-1834, 1991
A/Title: The nucleotide sequence and luteovirus-like nature of RNA 1 of an aphid non-trans
A/Reference number: JQ1382; MUID:91341468; PMID:1875194
A/Accession: JQ1383
A/Molecule type: genomic RNA
A/Residues: 1-760 <DEM>
A/Cross-references: UNIPROT:Q84710; GB:L04573; NID:G294105; PIDN:AAA72300.1; PID:G294109
A/Experimental source: strain WSG
A/Note: 271-Thr, 275-Phe, 502-Gly, 503-Arg, 571-Met, 639-Lys, 660-Gln, 666-Ser, 686-Ile,
C/Comment: This protein is involved in post-translational processing of viral translation
C/Genetics:
A/Map position: segment RNAL

Query Match      77.4%; Score 41; DB 2; Length 760;
Best Local Similarity 70.0%; Pred. No. 12;
Matches      7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 RLQEDPPAGV 10
      .|.|.|.|.|
Db      325 KLREDPPKGV 334

RESULT 14
T32959
hypothetical protein C35B1.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C/Accession: T32959
R/Miller, N.; Stelljes, L.; Bradshaw, H.; Keppler, D.
submitted to the EMBL Data Library, February 1998
A/Description: The sequence of C. elegans cosmid C35B1.
A/Reference number: Z21255
A/Accession: T32959
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-205 <ML>
A/Cross-references: EMBL:AF045638; PIDN:AAC02561.1; GSPDB:GN00022; CESP:C35B1.1
A/Experimental source: strain Bristol N2; clone C35B1
C/Genetics:
A/Gene: CESP:C35B1.1
A/Map position: 4
A/Introns: 17/3; 30/3; 160/3
C/Superfamily: yeast ubiquitin-protein ligase UBC1

Query Match      73.6%; Score 39; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      4 EDPPAGV 10
      .|.|.|.|
Db      31 EDPPAGV 37

RESULT 15
D64835
ABC-type transport protein uup - Escherichia coli (strain K-12)
C/Species: Escherichia coli
C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 16-Aug-2004
C/Accession: D64835
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: D64835
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-635 <BLAT>
A/Cross-references: UNIPROT:P43672; GB:AE000197; GB:U00096; NID:g1787180; PIDN:AAC74035.
A/Experimental source: strain K-12, substrain MG1655
C/Genetics:
A/Gene: uup
C/Superfamily: ATP-binding cassette homology
C/Keywords: ATP; nucleotide binding; P-loop; transport protein

```

F19-229/Domain: ATP-binding cassette homology <ABC1>  
 F36-43/Region: nucleotide-binding motif A (P-loop)  
 F335-513/Domain: ATP-binding cassette homology <ABC2>  
 F352-359/Region: nucleotide-binding motif A (P-loop)

Query Match 73.6%; Score 39; DB 2; Length 635;  
 Best Local Similarity 70.0%; Pred. No. 23;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10  
 |||:||||  
 Db 70 RLQDDPRNV 79

Search completed: November 18, 2005, 01:09:26  
 Job time : 11.9925 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:04:26 ; Search time 44.4086 Seconds  
(without alignments)  
115.311 Million cell updates/sec

Title: US-10-006-177-2

Perfect score: 53

Sequence: 1 RLQEDPPAGV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	100.0	124	Q95L52	Q95L52 bos taurus
2	53	100.0	124	Q9N1X5	Q9N1X5 bos taurus
3	53	100.0	126	Q9N1X6	Q9N1X6 equus caball
4	53	100.0	126	Q9N1X7	Q9N1X7 sus scrofa
5	53	100.0	152	1 UBCA_HUMAN	P49459 homo sapien
6	53	100.0	152	1 UBCA_MOUSE	Q9z255 mus musculu
7	53	100.0	152	2 ORJHF8	Q8jhf8 fundulus he
8	53	100.0	152	2 Q8GL05	Q8g105 xenopus tro
9	53	100.0	152	2 Q8PB14	Q8pb14 xenopus lae
10	53	100.0	152	2 Q8Y216	Q8y216 pagrus majo
11	53	100.0	152	2 Q7ZZA8	Q7zza8 brachydanio
12	53	100.0	152	2 Q803M4	Q803m4 brachydanio
13	53	100.0	152	2 Q8W6F3	Q8w6f3 gallus gall
14	53	100.0	162	2 Q8AXR9	Q8axr9 rattus norv
15	50	94.3	192	1 UBC1_CABEL	P52478 caenorhabdi
16	49	92.5	151	1 UBC6_DROME	P25153 drosophila
17	49	92.5	152	1 UBC1_ARATH	P25865 arabidopsis
18	49	92.5	152	1 UBC2_ARATH	P42745 arabidopsis
19	49	92.5	152	1 UBC2_HUMAN	P63146 homo sapien
20	49	92.5	152	1 UBC2_MOUSE	P63147 mus musculu
21	49	92.5	152	1 UBC2_RABIT	P63148 oryctolagus
22	49	92.5	152	1 UBC2_RAT	P63149 rattus norv
23	49	92.5	152	1 UBC2_WHEAT	P25866 triticum ae
24	49	92.5	152	2 Q8SBC1	Q8sbc1 oryza sativ
25	49	92.5	152	2 Q7XIC8	Q7xic8 oryza sativ
26	49	92.5	152	2 Q9AVN9	Q9avn9 nicotiana t
27	49	92.5	152	2 Q9AVP0	Q9avp0 nicotiana t
28	49	92.5	152	2 Q9M4R0	Q9m4r0 avicennia m
29	49	92.5	152	2 Q6DDC8	Q6ddc8 xenopus tro
30	49	92.5	152	2 Q6DGG1	Q6dgg1 brachydanio
31	49	92.5	152	2 Q6GR44	Q6gr44 xenopus lae

32 49 92.5 153 2 Q7PFQ9 Q7pfq9 anopheles g  
33 49 92.5 157 2 Q7QFT1 Q7qft1 anopheles g  
34 48 90.6 151 1 UBC2\_SCHPO P23566 schizosacch  
35 47 88.7 152 1 UBC2\_MEISA P35130 medicago sa  
36 45 84.9 151 1 UBC2\_NEUCR P52493 neurospora  
37 45 84.9 151 2 P787I7 P787i7 nectria hae  
38 45 84.9 151 2 Q96UP5 Q96up5 emericella  
39 45 84.9 151 2 Q7RVK4 Q7rvk4 neurospora  
40 45 84.9 179 1 UBC2\_CANAL Q74201 candida alb  
41 44 83.0 146 2 Q94AD2 Q94ad2 arabidopsis  
42 44 83.0 146 2 Q8H1P1 Q8h1p1 arabidopsis  
43 44 83.0 150 1 UBC3\_ARATH P42746 arabidopsis  
44 43 81.1 151 2 Q6C0S3 Q6c0s3 yarrowia li  
45 42 79.2 154 2 Q86F26 Q86f26 schistosoma

#### ALIGNMENTS

##### RESULT 1

Q95L52 ID Q95L52 PRELIMINARY; PRT; 124 AA.  
AC Q95L52;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE RAD6-like protein HR6A (Fragment).  
GN Name=BHR6A;  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Polounienko A., Blecher S.R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
CC diphosphate + protein N-ubiquityllysine.  
CC -!- PATHWAY: Ubiquitin conjugation; second step.  
CC -!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
DR EMBL; AF419316; AAL14998.1; .  
DR HSP; P23567; IJAS.  
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000608; UBQ\_conjugat.  
DR Pfam; PF00179; UQ\_con; 1\_  
DR ProDom; PD000461; UBQ\_conjugat; 1.  
DR SMART; SMO0212; UBCC; 1.  
DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
DR PROSITE; PS0127; UBIQUITIN\_CONJUGAT\_2; 1.  
KW Ligase; Ubl conjugation pathway.  
FT NON\_TER 1 1  
FT NON\_TER 124 124  
SQ SEQUENCE 124 AA; 13879 MW; 1ADB62B015E3C6E9 CRC64;

Query Match 100.0%; Score 53; DB 2; Length 124;

Best Local Similarity 100.0%; Pred. No. 0.042;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10

Db 5 RLQEDPPAGV 14

##### RESULT 2

Q9N1X5 ID Q9N1X5 PRELIMINARY; PRT; 124 AA.

AC Q9N1X5;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

DE RAD6 homolog (Fragment).
GN Name=BHR6A;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP
RQ
SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Poloumienko A., Blecher S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.
DR HSSP; P23567; 1JAS.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR EMBL; AF132075; AAF36530.1; -.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000608; UBQ_conjugat.
DR Pfam; PF00179; UBQ_con; 1.
DR ProDom; PD000461; UBQ_conjugat; 1.
DR SMART; SM00212; UBQC_1.
DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubl conjugation pathway.
FT NON TER 1
FT NON TER 124
SQ SEQUENCE 124 AA; 13878 MW; 90F3C03A3561C8E9 CRC64;

Query Match 100.0%; Score 53; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10
DB 5 RLQEDPPAGV 14

RESULT 3
Q9NIX6 PRELIMINARY; PRT; 126 AA.
AC Q9NIX6
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RAD6 homolog (Fragment).
GN Name=BHR6A;
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP
RQ
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Poloumienko A., Blecher S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.
DR HSSP; P23567; 1JAS.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000608; UBQ_conjugat.
DR Pfam; PF00179; UBQ_con; 1.
DR ProDom; PD000461; UBQ_conjugat; 1.
DR SMART; SM00212; UBQC_1.
DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubl conjugation pathway.

```

```

FT NON TER 1
FT NON TER 126
SQ SEQUENCE 126 AA; 14148 MW; FAC0F3A7A1E2184 CRC64;

Query Match 100.0%; Score 53; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10
DB 6 RLQEDPPAGV 15

RESULT 4
Q9NIX7 PRELIMINARY; PRT; 126 AA.
AC Q9NIX7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RAD6 homolog (Fragment).
GN Name=BHR6A;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP
RQ
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Poloumienko A., Blecher S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.
DR HSSP; P23567; 1JAS.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000608; UBQ_conjugat.
DR Pfam; PF00179; UBQ_con; 1.
DR ProDom; PD000461; UBQ_conjugat; 1.
DR SMART; SM00212; UBQC_1.
DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubl conjugation pathway.
FT NON TER 1
FT NON TER 126
SQ SEQUENCE 126 AA; 14132 MW; 1ACD93A6C2B4338D CRC64;

Query Match 100.0%; Score 53; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10
DB 6 RLQEDPPAGV 15

RESULT 5
UBCA_HUMAN STANDARD; PRT; 152 AA.
ID UBCA_HUMAN
AC P49459; Q96FX4;
DT 01-FEB-1996 (Rel. 33, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ubiquitin-conjugating enzyme E2 A (EC 6.3.2.19) (Ubiquitin-protein
DE ligase A) (Ubiquitin carrier protein A) (HR6A) (hHR6A).
GN Name=UBE2A; Synonyms=RAD6A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

RN RP SEQUENCE FROM N.A.  
 RX MEDLINE=92020951; PubMed=1717990;  
 RA Koken M.H.M., Reynolds P., Jaspers-Dekker I., Prakash L., Prakash S.,  
 RA Bootsma D., Hoeijmakers J.H.J.;  
 RT "Structural and functional conservation of two human homologs of the  
 RT yeast DNA repair gene RAD6";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8865-8869(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshlyuk S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Catalyzes the covalent attachment of ubiquitin to other  
 CC proteins. Required for postreplication repair of UV-damaged DNA.  
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC -!- SUBUNIT: Ubiquitin conjugation; second step.  
 CC -!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M74524; AAA35981.1; -;  
 CC EMBL; BC010175; AAH10175.1; -;  
 CC PIR; A41222; A41222.  
 CC HSSP; P25865; 2AAK.  
 CC Genew; HGNC:12472; UBE2A.  
 CC MIW; 312180; -;  
 CC GO; GO:0004840; F:ubiquitin conjugating enzyme activity; TAS.  
 CC GO; GO:0006301; P:postreplication repair; TAS.  
 CC GO; GO:0006511; P:ubiquitin-dependent protein catabolism; TAS.  
 CC InterPro; IPR000608; UBQ conjugat.  
 CC Pfam; PF00179; UBQ con; 1.  
 CC ProDom; PD000461; UBQ conjugat; 1.  
 CC SMART; SM00212; UBQs; 1.  
 CC PROSITE; PS00189; UBIQUITIN CONJUGAT\_1; 1.  
 CC PROSITE; PS00127; UBIQUITIN CONJUGAT\_2; 1.  
 CC DNA repair; Ligase; Multigene family; Nuclear protein;  
 KW Ub1 conjugation pathway.  
 FT ACT\_SITE 88  
 FT 88 Glycyl thioester intermediate (By  
 FT 49  
 FT 49 E -> G (in Ref. 1).  
 FT CONFLICT 152 AA; 17315 MW; 0AAEB5B770B47E2 CRC64;  
 SQ SEQUENCE 100.0%; Score 53; DB 1; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.053;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10  
 DB 15 RLQEDPPAGV 24  
 RESULT 6  
 UBCA MOUSE  
 ID UBCA MOUSE STANDARD; PRT; 152 AA.  
 AC Q94255;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Ubiquitin-conjugating enzyme E2 A (EC 6.3.2.19) (Ubiquitin-protein  
 DE ligase A) (Ubiquitin carrier protein A) (HR6A) (mHR6A).  
 GN Name=Ube2a; Synonyms=Rad6a;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BCBA; TISSUE=Brain;  
 RA Roest H.P., van Klaveren J., Koken M.H.M., Vermeij M.,  
 RA van Cappellen W.A., Baarends W.M., Hoogerbrugge J.W., Bootsma D.,  
 RA Hoeijmakers J.H.J., Grootegoed J.A., de Wit J.;  
 RA "Isolation of mHR6A, a gene highly homologous to the male fertility  
 RT gene mHR6B";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Skeletal muscle;  
 RA Kwon Y.T., Varshavsky A.;  
 RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshlyuk S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Catalyzes the covalent attachment of ubiquitin to other  
 CC proteins. Required for postreplication repair of UV-damaged DNA.  
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC -!- SUBUNIT: Ubiquitin conjugation; second step.  
 CC -!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M74524; AAA35981.1; -;  
 CC EMBL; BC010175; AAH10175.1; -;  
 CC PIR; A41222; A41222.  
 CC HSSP; P25865; 2AAK.  
 CC Genew; HGNC:12472; UBE2A.  
 CC MIW; 312180; -;  
 CC GO; GO:0004840; F:ubiquitin conjugating enzyme activity; TAS.  
 CC GO; GO:0006301; P:postreplication repair; TAS.  
 CC GO; GO:0006511; P:ubiquitin-dependent protein catabolism; TAS.  
 CC InterPro; IPR000608; UBQ conjugat.  
 CC Pfam; PF00179; UBQ con; 1.  
 CC ProDom; PD000461; UBQ conjugat; 1.  
 CC SMART; SM00212; UBQs; 1.  
 CC PROSITE; PS00189; UBIQUITIN CONJUGAT\_1; 1.  
 CC PROSITE; PS00127; UBIQUITIN CONJUGAT\_2; 1.  
 CC DNA repair; Ligase; Multigene family; Nuclear protein;  
 KW Ub1 conjugation pathway.  
 FT ACT\_SITE 88  
 FT 88 Glycyl thioester intermediate (By  
 FT 49  
 FT 49 E -> G (in Ref. 1).  
 FT CONFLICT 152 AA; 17315 MW; 0AAEB5B770B47E2 CRC64;  
 SQ SEQUENCE 100.0%; Score 53; DB 1; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.053;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q6GL05	PRELIMINARY;	PRT;	152 AA.
Q6GL05	AC		
Q6JUL-2004	(TREMBLrel. 27, Created)		
05-JUL-2004	(TREMBLrel. 27, Last sequence update)		
05-JUL-2004	(TREMBLrel. 27, Last annotation update)		
MGCE9378	protein.		
DE			
GN	Name=MGCE9378;		
OS	Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;		
OC	Xenopodinae; Xenopus.		
OX	NCBI_TaxID=8364;		
11			
RN	SEQUENCE FROM N.A.		
RP	TISSUE=Embryo;		
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Bhat N.K.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.J., Hsieh F.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Scheetz T.E.,		
RA	Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Prange C.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullaly S.J.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Murthy D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blatesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Matra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
12			
RN	SEQUENCE FROM N.A.		
RP	TISSUE=Embryo;		
RC	Klein S., Gerhard D.S.;		
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.		
RA	-!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +		
CC	diphosphate + protein N-ubiquityllysine.		
CC	-!- PATHWAY: Ubiquitin conjugation; second step.		
CC	-!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.		
CC	EMBL: BC074715; AAH74715.1; -		
DR	GO: GO:0016874; F.ligase activity; IEA.		
DR	GO: GO:0004840; F.ubiquitin conjugating enzyme activity; IEA.		
DR	GO: GO:0004512; F.ubiquitin cycle; IEA.		
DR	InterPro: IPR000608; UBQ_conjugat.		
DR	Pfam: PF00179; UQ_con; 1.		
DR	ProDom: PD000461; UBQ_conjugat; 1.		
DR	SMART: SM00212; UBCC_1.		
DR	PROSITE: PS00183; UBIQUITIN_CONUGAT_1; 1.		
DR	PROSITE: PS0127; UBIQUITIN_CONUGAT_2; 1.		
KW	Ligase; Ub1 conjugation pathway.		
SW	SSQUENCE 152 AA; 17315 MW; 0AAB5B7770E47E2 CRC64;		
Query Match	100.0%;	Score 53;	DB 2; Length 152;
Best Local Similarity	100.0%;	Pred. No. 0.053;	
Matches 10;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0
QY	1 RLQEDPPAGV 10		
DB	15 RLQEDPPAGV 24		
RESULT 9			
Q6PB14	PRELIMINARY;	PRT;	152 AA.
ID	Q6PB14		
AC	Q6PB14;		
DT	05-JUL-2004 (TREMBLrel. 27, Created)		
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)		

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE MGCS8540 protein.  
 GN Name=MGCS8540;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Griinwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzyzanski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RL Dev. Dyn. 225:384-391 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Klein S., Strausberg R.;  
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC -!- PATHWAY: Ubiquitin conjugation; second step.  
 CC -!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
 DR EMBL; BC059970; AAHS9970.1; -.  
 DR HSSP; P06104; IAYZ.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.  
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000608; UBQ\_conjugat.  
 DR Pfam; PF00179; UBQ\_con; 1.  
 DR ProDom; PD000461; UBQ\_conjugat; 1.  
 DR SMART; SM00212; UBCC; 1.  
 DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
 DR PROSITE; PS0127; UBIQUITIN\_CONJUGAT\_2; 1.  
 KW Ligase; Ubi conjugation pathway.  
 SQ SEQUENCE 152 AA; 17315 MW; 0AAEB5B7770B47E2 CRC64;  
 Query Match 100.0%; Score 53; DB 2; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.053;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 RLQEDPPAGV 10  
 Db 15 RLQEDPPAGV 24  
 |||||  
 |||||

RESULT 10  
 Q5Y216  
 ID Q6Y216 PRELIMINARY; PRT; 152 AA.  
 AC Q6Y216;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Ubiquitin-conjugating enzyme E2A.  
 OS Pagrus major (Red sea bream) (Chrysophrys major).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 OC Sparidae; Pagrus.  
 OX NCBI\_TaxID=143350;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RA Chen S.L., Xu M.Y., Hu S.L., Li L.;  
 RT "Analysis of immune-relevant genes expressed in red sea bream  
 RT spleen";  
 RL Aquaculture 240:115-130 (2004).  
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC -!- PATHWAY: Ubiquitin conjugation; second step.  
 CC -!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
 DR EMBL; AY190722; AAP20197.1; -.  
 DR HSSP; P06104; IAYZ.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.  
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000608; UBQ\_conjugat.  
 DR Pfam; PF00179; UBQ\_con; 1.  
 DR ProDom; PD000461; UBQ\_conjugat; 1.  
 DR SMART; SM00212; UBCC; 1.  
 DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
 DR PROSITE; PS0127; UBIQUITIN\_CONJUGAT\_2; 1.  
 KW Ligase; Ubi conjugation pathway.  
 SQ SEQUENCE 152 AA; 17297 MW; 8AD3ED62FA45290C CRC64;  
 Query Match 100.0%; Score 53; DB 2; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.053;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 RLQEDPPAGV 10  
 Db 15 RLQEDPPAGV 24  
 |||||  
 |||||

RESULT 11  
 Q7ZZA8  
 ID Q7ZZA8 PRELIMINARY; PRT; 152 AA.  
 AC Q7ZZA8;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE SI:B246J2.4 (Novel ubiquitin-conjugating enzyme) (Ubiquitin-  
 DE conjugating enzyme E2A) (RAD6 homolog).  
 GN Name=ube2a; Synonyms=SI:B246J2.4;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Skuce C.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Griinwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzyzanski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RL Dev. Dyn. 225:384-391 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Klein S., Strausberg R.;  
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC -!- PATHWAY: Ubiquitin conjugation; second step.  
 CC -!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
 DR EMBL; BC059970; AAHS9970.1; -.  
 DR HSSP; P06104; IAYZ.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.  
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000608; UBQ\_conjugat.  
 DR Pfam; PF00179; UBQ\_con; 1.  
 DR ProDom; PD000461; UBQ\_conjugat; 1.  
 DR SMART; SM00212; UBCC; 1.  
 DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
 DR PROSITE; PS0127; UBIQUITIN\_CONJUGAT\_2; 1.  
 KW Ligase; Ubi conjugation pathway.  
 SQ SEQUENCE 152 AA; 17315 MW; 0AAEB5B7770B47E2 CRC64;  
 Query Match 100.0%; Score 53; DB 2; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.053;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 RLQEDPPAGV 10  
 Db 15 RLQEDPPAGV 24  
 |||||  
 |||||

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Carninci P., Mullaly S.J.,  
RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalon D.K., Wozny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalon D.K., Wozny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
CC diphosphate + protein N-ubiquityllysine.  
CC -1- PATHWAY: Ubiquitin conjugation; second step.  
CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
DR EMBL; AL590151; CAB68063.1; -.  
DR EMBL; BC053256; AAH53256.1; -.  
DR HSP; P23567; IJAS.  
DR ZFIN; ZDB-GENE-030616-72; ube2a.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000608; UBQ\_conjugat.  
DR Pfam; PF00179; UQ\_con; 1.  
DR ProDom; PD000461; UBQ\_conjugat; 1.  
DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
DR PROSITE; PS0127; UBIQUITIN\_CONJUGAT\_2; 1.  
KW Ligase; Ubl conjugation pathway.  
SQ SEQUENCE 152 AA; 17315 MW; 0AAEB5B7707B47E2 CRC64;

Query Match 100.0%; Score 53; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.053;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
| | | | | | | | | |  
Db 15 RLQEDPPAGV 24

RESULT 12  
Q803M4 PRELIMINARY; PRT; 152 AA.  
AC Q803M4;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Ubiquitin-conjugating enzyme E2B (RAD6 homolog).  
GN ORFNames=rgc:55512;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB; TISSUE=whole body;  
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant P., Prange C.,  
RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Mullaly S.J.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Carninci P., Prange C.,  
RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalon D.K., Wozny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalon D.K., Wozny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB; TISSUE=Whole body;  
RA Strausberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
CC diphosphate + protein N-ubiquityllysine.  
CC -1- PATHWAY: Ubiquitin conjugation; second step.  
CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
DR EMBL; BC044416; AAH44416.1; -.  
DR HSP; P23567; IJAS.  
DR ZFIN; ZDB-GENE-030131-4195; zgc:55512.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000608; UBQ\_conjugat.  
DR Pfam; PF00179; UQ\_con; 1.  
DR ProDom; PD000461; UBQ\_conjugat; 1.  
DR SMART; SM00212; UBCc; 1.  
DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
DR PROSITE; PS0127; UBIQUITIN\_CONJUGAT\_2; 1.  
KW Ligase; Ubl conjugation pathway.  
SQ SEQUENCE 152 AA; 17285 MW; 9A2D0D8CF4A5D7FC CRC64;

Query Match 100.0%; Score 53; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.053;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
| | | | | | | | | |  
Db 15 RLQEDPPAGV 24

RESULT 13  
Q9W6F3 PRELIMINARY; PRT; 152 AA.  
AC Q9W6F3;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Ubiquitin-conjugating enzyme.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Sofer L., Burnside J.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
CC diphosphate + protein N-ubiquityllysine.  
CC -1- PATHWAY: Ubiquitin conjugation; second step.  
CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
DR EMBL; AF120212; AA031646.1; -.  
DR HSP; P23567; IJAS.  
DR GO; GO:0016874; F:ligase activity; IEA.

DR GO: 0004840; E-ubiquitin conjugating enzyme activity; IEA.  
 DR GO: 0006512; P-ubiquitin cycle; IEA.  
 DR InterPro: IPR00608; UBQ\_conjugat.  
 DR Pfam: PF00179; UQ\_con; 1.  
 DR ProDom: PD000461; UBQ\_conjugat; 1.  
 DR PROSITE: PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
 DR PROSITE: PS00127; UBIQUITIN\_CONJUGAT\_2; 1.  
 KW Ligase: Ubl conjugation pathway.  
 SQ SEQUENCE 152 AA; 17315 MW; 0AAEB5B7770E47E2 CRC64;

Query Match 100.0%; Score 53; DB 2; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.053;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10  
 DB 15 RLQEDPPAGV 24  
 |||||

RESULT 14

Q6AXR9 PRELIMINARY; PRT; 162 AA.  
 AC Q6AXR9;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahey J., Helton E., Kettner M., Madan A.C., Rodrigues S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 Jones S.J., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Director MGC Project;  
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC -1- PATHWAY: Ubiquitin conjugation; second step.  
 CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
 CC EMBL; BC079353; AAH79353.1; -  
 DR GO: 0004840; E-ubiquitin conjugating enzyme activity; IEA.  
 DR GO: 0006512; P-ubiquitin cycle; IEA.  
 DR InterPro: IPR00608; UBQ\_conjugat.  
 DR Pfam: PF00179; UQ\_con; 1.  
 DR ProDom: PD000461; UBQ\_conjugat; 1.  
 DR SMART: SM00212; UBCc; 1.  
 DR PROSITE: PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
 DR PROSITE: PS00127; UBIQUITIN\_CONJUGAT\_2; 1.

KW Hypothetical protein; Ligase; Ubl conjugation pathway.  
 SQ SEQUENCE 162 AA; 18339 MW; FED3471058FASD37 CRC64;

Query Match 100.0%; Score 53; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 0.057;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10  
 DB 15 RLQEDPPAGV 24  
 |||||

RESULT 15

UBC1\_CABEL STANDARD; PRT; 192 AA.  
 AC P52478; O45062;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Ubiquitin-conjugating enzyme E2 1 (EC 6.3.2.19) (Ubiquitin-protein  
 DE ligase 1) (Ubiquitin carrier protein 1).  
 GN Name=ubc-1; ORFNames=C35B1.1;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peleoderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=96027757; PubMed=7546294;  
 RA Leggett D.S., Jones D., Candido E.P.M.;  
 RT "Caenorhabditis elegans UBC-1, a ubiquitin-conjugating enzyme  
 RT homologous to yeast RAD6/UBC2, contains a novel carboxy-terminal  
 RT extension that is conserved in nematodes.";  
 RL DNA Cell Biol. 14:883-891 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA The C. elegans sequencing consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology";  
 RL Science 282:2012-2018 (1998).  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Catalyzes the covalent attachment of ubiquitin to other  
 CC proteins.  
 CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC -1- PATHWAY: Ubiquitin conjugation; second step.  
 CC -1- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U08139; AAA83388.1; -  
 CC PDB; AF045638; AAC02561.2; -  
 DR PDB; 1Q34; X-ray; A/B/C=1-163.  
 DR InterAct; P52478; -  
 DR WormBase; WBGene00006701; ubc-1.  
 DR WormPepe; C35B1.1; CE27822.  
 DR InterPro; IPR000608; UBQ\_conjugat.  
 DR Pfam; PF00179; UQ\_con; 1-  
 DR ProDom; PD000461; UBQ\_conjugat; 1.  
 DR SMART; SM00212; UBCc; 1.  
 DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.

DR PROSITE; PS50127; UBIQUITIN CONJUGAT 2; 1.  
KW 3D-structure; Ligase; MultiGene family; Ubl conjugation pathway.  
FT ACT\_SITE 88 Glycyl thioester intermediate (By  
FT similarity).  
SQ SEQUENCE 192 AA; 21513 MW; 7CF26B8F896EF33D CRC64;

Query Match 94.3%; Score 50; DB 1; Length 192;  
Best Local Similarity 90.0%; Pred. No. 0.26;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLQEDPPAGV 10  
Db 15 KLQEDPPAGV 24

Search completed: November 18, 2005, 01:07:53  
Job time : 48.4086 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 00:48:52 ; Search time 15.1613 Seconds  
(without alignments)  
49.237 Million cell updates/sec

Title: US-10-006-177-2

Perfect score: 53

Sequence: 1 RLQEDPPAGV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	100.0	57	4	US-09-513-999C-6307
2	53	100.0	126	4	US-09-949-016-7772
3	53	100.0	130	4	US-09-621-976-4511
4	53	100.0	152	1	US-08-318-947A-8
5	53	100.0	152	2	US-08-795-303-8
6	53	100.0	152	4	US-09-216-430C-13
7	53	100.0	152	4	US-09-538-092-1155
8	49	92.5	151	1	US-08-318-947A-9
9	49	92.5	151	2	US-08-795-303-9
10	49	92.5	152	1	US-08-318-947A-6
11	49	92.5	152	1	US-08-318-947A-7
12	49	92.5	152	2	US-08-533-298-2
13	49	92.5	152	2	US-08-795-303-6
14	49	92.5	152	2	US-08-795-303-7
15	49	92.5	152	2	US-08-247-904B-16
16	49	92.5	152	3	US-08-767-942A-17
17	49	92.5	152	3	US-09-358-580-8
18	49	92.5	152	3	US-09-358-580-10
19	49	92.5	152	4	US-09-216-430C-12
20	49	92.5	152	4	US-09-538-092-1000
21	49	92.5	257	4	US-09-949-016-8138
22	48	90.6	151	1	US-08-318-947A-10
23	48	90.6	151	2	US-08-795-303-10
24	44	83.0	152	3	US-09-358-580-12
25	39	73.6	636	4	US-09-489-039A-12090
26	39	73.6	722	4	US-09-252-991A-26839
27	38	71.7	695	4	US-09-252-991A-22910

## ALIGNMENTS

### RESULT 1

US-09-513-999C-6307  
; Sequence 6307, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59. US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513.999C  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6307  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-513-999C-6307

Query Match 100.0%; Score 53; DB 4; Length 57;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
| | | | |  
Db 15 RLQEDPPAGV 24

### RESULT 2

US-09-949-016-7772  
; Sequence 7772, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7772
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7772

Query Match      100.0%; Score 53; DB 4; Length 126;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RLQEDPPAGV 10
Db      55 RLQEDPPAGV 64

RESULT 3
US-09-621-976-4511
; Sequence 4511, Application US/09621976
; Patent No. 6839063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSER.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4511
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 5
; OTHER INFORMATION: Xaa = * , Cys, Asp, Glu, Gly, Tyr
; NAME/KEY: UNSURE
; LOCATION: 14
; OTHER INFORMATION: Xaa = Ala, Gly
; NAME/KEY: UNSURE
; LOCATION: 7,16,17
; OTHER INFORMATION: Xaa = Ala, Gly, Pro, Arg
; NAME/KEY: UNSURE
; LOCATION: 11
; OTHER INFORMATION: Xaa = Ala, Phe, Ser, Val
; NAME/KEY: UNSURE
; LOCATION: 9
; OTHER INFORMATION: Xaa = Asp, Glu, Gly, His, Gln, Arg
; NAME/KEY: UNSURE
; LOCATION: 8
; OTHER INFORMATION: Xaa = Cys, Pro, Arg, Ser, Trp
; NAME/KEY: UNSURE
; LOCATION: 3,12
; OTHER INFORMATION: Xaa = His, Asn, Pro, Thr
; NAME/KEY: UNSURE
; LOCATION: 89
; OTHER INFORMATION: Xaa = His, Gln
; NAME/KEY: UNSURE
; LOCATION: 10
; OTHER INFORMATION: Xaa = His, Ile, Lys, Leu, Met, Asn, Gln
; NAME/KEY: UNSURE
; LOCATION: 15
; OTHER INFORMATION: Xaa = Pro, Arg, Ser, Trp
; NAME/KEY: UNSURE
; LOCATION: 90
; OTHER INFORMATION: Xaa = Pro, Ser
US-09-621-976-4511

Query Match      100.0%; Score 53; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 RLQEDPPAGV 10
Db      24 RLQEDPPAGV 33

RESULT 4
US-08-318-947A-8
; Sequence 8, Application US/08318947A
; Patent No. 5798245
; GENERAL INFORMATION:
; APPLICANT: Anderson, Paul J.
; APPLICANT: Tian, Qingsheng
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, NW Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,947A
; FILING DATE: 06-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/133,530
; FILING DATE: 07-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mack, Susan J.
; REGISTRATION NUMBER: 30,951
; REFERENCE/DOCKET NUMBER: A6462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-2920
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-318-947A-8

Query Match      100.0%; Score 53; DB 1; Length 152;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RLQEDPPAGV 10
Db      15 RLQEDPPAGV 24

RESULT 5
US-08-795-303-8
; Sequence 8, Application US/08795303
; Patent No. 5948656
; GENERAL INFORMATION:
; APPLICANT: Anderson, Paul J.
; APPLICANT: Tian, Qingsheng
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
```

STREET: 2100 Pennsylvania Avenue, NW Suite 800  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20037

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,303  
FILING DATE: 04-FEB-1997  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/318,947  
FILING DATE: 06-OCT-1994  
APPLICATION NUMBER: 08/133,530  
FILING DATE: 07-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mack, Susan J.  
REGISTRATION NUMBER: 30,951  
REFERENCE/DOCKET NUMBER: A6462  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-2920  
TELEX: 6491103

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-795-303-8

Query Match 100.0%; Score 53; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.039;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPAGV 24

RESULT 6  
US-09-216-430C-13  
Sequence 13, Application US/09216430C  
Patent No. 6734283

GENERAL INFORMATION:  
APPLICANT: Chau, Vincent  
TITLE OF INVENTION: Human Proteins Responsible for NEDD8 Activation and Conjugation  
FILE REFERENCE: 103576-127  
CURRENT APPLICATION NUMBER: US/09/216,430C

CURRENT FILING DATE: 1998-12-18  
PRIOR APPLICATION NUMBER: PCT/US98/27141  
PRIOR FILING DATE: 1998-12-18  
PRIOR APPLICATION NUMBER: US 60/068,209  
PRIOR FILING DATE: 1998-08-12  
NUMBER OF SEQ ID NOS: 38

SOFTWARE: PatentIn version 3.1  
SEQ ID NO 13  
LENGTH: 152  
TYPE: PRT

ORGANISM: Human  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1)..(152)  
OTHER INFORMATION: UBC2a

US-09-216-430C-13

Query Match 100.0%; Score 53; DB 4; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.039;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPAGV 24

RESULT 7

US-09-538-092-1155  
Sequence 1155, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:

APPLICANT: Giot, Loic  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CuratSeqFormatter Version 0.9  
SEQ ID NO 1155  
LENGTH: 152  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)..(0)

OTHER INFORMATION: Polypeptide Accession Number P49459  
US-09-538-092-1155

Query Match 100.0%; Score 53; DB 4; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.039;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPAGV 24

RESULT 8

US-08-318-947A-9  
Sequence 9, Application US/08318947A  
Patent No. 5798245

GENERAL INFORMATION:  
APPLICANT: Anderson, Paul J.  
APPLICANT: Tian, Qingsheng  
TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED  
TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas  
STREET: 2100 Pennsylvania Avenue, NW Suite 800  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,947A  
FILING DATE: 06-OCT-1994  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/133,530  
FILING DATE: 07-OCT-1993  
ATTORNEY/AGENT INFORMATION:

```

; NAME: Mack, Susan J.
; REGISTRATION NUMBER: 30,951
; REFERENCE/DOCKET NUMBER: A6462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-2920
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-318-947A-9

Query Match 92.5%; Score 49; DB 1; Length 151;
Best Local Similarity 90.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10
Db 15 RLQEDPPTGV 24

RESULT 9
US-08-795-303-9
; Sequence 9, Application US/08795303
; Patent No. 5948656
; GENERAL INFORMATION:
; APPLICANT: Anderson, Paul J.
; APPLICANT: Tian, Qingsheng
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, NW Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,303
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/318,947
; FILING DATE: 06-OCT-1994
; APPLICATION NUMBER: 08/133,530
; FILING DATE: 07-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mack, Susan J.
; REGISTRATION NUMBER: 30,951
; REFERENCE/DOCKET NUMBER: A6462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-2920
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-795-303-9

Query Match 92.5%; Score 49; DB 1; Length 151;
Best Local Similarity 90.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10
Db 15 RLQEDPPTGV 24

RESULT 10
US-08-318-947A-6
; Sequence 6, Application US/08318947A
; Patent No. 5798245
; GENERAL INFORMATION:
; APPLICANT: Anderson, Paul J.
; APPLICANT: Tian, Qingsheng
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, NW Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,947A
; FILING DATE: 06-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/133,530
; FILING DATE: 07-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mack, Susan J.
; REGISTRATION NUMBER: 30,951
; REFERENCE/DOCKET NUMBER: A6462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-2920
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-318-947A-6

Query Match 92.5%; Score 49; DB 1; Length 152;
Best Local Similarity 90.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10
Db 15 RLQEDPPTGV 24

RESULT 11
US-08-318-947A-7
; Sequence 7, Application US/08318947A
; Patent No. 5798245
; GENERAL INFORMATION:
; APPLICANT: Anderson, Paul J.
; APPLICANT: Tian, Qingsheng
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
; US-08-318-947A-7
```

```

; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zimm, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, NW Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,947A
; FILING DATE: 06-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/133,530
; FILING DATE: 07-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mack, Susan J.
; REGISTRATION NUMBER: 30,951
; REFERENCE/DOCKET NUMBER: A6462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-2920
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-318-947A-7

```

Query Match 92.5%; Score 49; DB 1; Length 152;  
Best Local Similarity 90.0%; Pred. No. 0.2;  
Matches 9; Conservative 0; Mismatches 1; Indels

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPVGV 24

```

RESULT 12
US-08-533-298-2
; Sequence 2, Application US/08533298
; Patent No. 5851791
; GENERAL INFORMATION:
; APPLICANT: Vierastra, Richard D
; APPLICANT: Gosink, Mark M
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme (E2) Fusion
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme (E2) Fusion
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles and Brady
; STREET: 1 South Pinckney Street - Suite 600
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/533,298
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:

```

```

; APPLICATION NUMBER: 08/070,157
; FILING DATE: 28-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-533-298-2

Query Match          92.5%; Score 49; DB 2; Length 152;
Best Local Similarity 80.0%; Pred. NO. 0.2;
Matches 8; Conservative 0; Indels

```

Qy	1	RLQEDPPAGV	10
		:	
Db	15	RLQODPPAGI	24

RESULT 13  
US-08-795-303-6  
; Sequence 6, Application US/08795303  
; Patent No. 5948656  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Paul J.  
; APPLICANT: Tian, Qingsheng  
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED  
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas  
; STREET: 2100 Pennsylvania Avenue, NW Suite 800  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/795,303  
; FILING DATE: 04-FEB-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/318,947  
; FILING DATE: 06-OCT-1994  
; APPLICATION NUMBER: 08/133,530  
; FILING DATE: 07-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mack, Susan J.  
; REGISTRATION NUMBER: 30,951  
; REFERENCE/DOCKET NUMBER: A6462  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)293-7060  
; TELEFAX: (202)293-2920  
; TELEX: 6491103  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 152 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-795-303-6

Query Match 92.5%; Score 49; DB 2; Length 152;  
Best Local Similarity 90.0%; Pred. No. 0.2;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
| | | | | | | |  
Db 15 RLQEDPPVGV 24

## RESULT 14

US-08-795-303-7  
; Sequence 7, Application US/08795303  
; Patent No. 5948656  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Paul J.  
; APPLICANT: Tian, Qingheng  
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED  
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sughruue, Mion, Zinn, Macpeak & Seas  
; STREET: 2100 Pennsylvania Avenue, NW Suite 800  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20037

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/08/795.303  
; APPLICATION NUMBER: 08/318,947  
; FILING DATE: 04-FEB-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/133,530  
; FILING DATE: 06-OCT-1994  
; APPLICATION NUMBER: 08/133,530  
; FILING DATE: 07-OCT-1993

ATTORNEY/AGENT INFORMATION:  
; NAME: Mack, Susan J.  
; REGISTRATION NUMBER: 30,951  
; REFERENCE/DOCKET NUMBER: A6462  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)293-7060  
; TELEFAX: (202)293-2920  
; TELEX: 6491103

INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 152 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-795-303-7

Query Match 92.5%; Score 49; DB 2; Length 152;  
Best Local Similarity 90.0%; Pred. No. 0.2;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
| | | | | | | |  
Db 15 RLQEDPPVGV 24

## RESULT 15

US-08-247-904B-16  
; Sequence 16, Application US/08247904B  
; Patent No. 5981699  
; GENERAL INFORMATION:  
; APPLICANT: Rolfe, Mark  
; APPLICANT: Eckstein, Jens W.

; APPLICANT: Draetta, Giulio  
; TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley, Hoag & Eliot  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/247,904B  
; FILING DATE: 23-MAY-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MIV-029.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 152 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-247-904B-16

Query Match 92.5%; Score 49; DB 2; Length 152;  
Best Local Similarity 90.0%; Pred. No. 0.2;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
| | | | | | | |  
Db 15 RLQEDPPVGV 24

Search completed: November 18, 2005, 09:35:53  
Job time : 16.1613 secs

GenCore version 5.1.6  
Copyright (C) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 01:08:37 ; Search time 49.2473 Seconds  
(without alignments)  
84.961 Million cell updates/sec

Title: US-10-006-177-2

Perfect score: 53

Sequence: 1 RLQEDPPAGV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*

17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*

18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*

19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*

20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*

21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	100.0	10	14	US-10-006-177-2
2	53	100.0	89	16	Sequence 2, Appli
3	53	100.0	122	17	Sequence 234519,
4	53	100.0	126	17	Sequence 15, Appl
5	53	100.0	126	17	Sequence 18529, A
6	53	100.0	126	17	Sequence 18530, A
7	49	92.5	58	16	Sequence 145487,
8	49	92.5	69	16	Sequence 217253,
9	49	92.5	73	15	Sequence 145487,
10	49	92.5	77	15	Sequence 282590,
11	49	92.5	82	15	Sequence 194713,

12	49	92.5	83	15	US-10-424-599-169060	Sequence 169060,
13	49	92.5	93	16	US-10-425-115-218996	Sequence 218996,
14	49	92.5	115	16	US-10-425-115-225960	Sequence 225960,
15	49	92.5	121	16	US-10-425-115-217252	Sequence 217252,
16	49	92.5	141	16	US-10-425-115-218993	Sequence 218993,
17	49	92.5	151	20	US-11-097-143-3093	Sequence 3093, Ap
18	49	92.5	152	15	US-10-424-599-158042	Sequence 158042,
19	49	92.5	152	15	US-10-424-599-194711	Sequence 194711,
20	49	92.5	152	15	US-10-424-599-280313	Sequence 280313,
21	49	92.5	152	15	US-10-424-599-280314	Sequence 280314,
22	49	92.5	152	16	US-10-437-963-148925	Sequence 148925,
23	49	92.5	152	16	US-10-767-701-46376	Sequence 46376, A
24	49	92.5	152	16	US-10-767-701-46377	Sequence 46377, A
25	49	92.5	152	16	US-10-425-115-217248	Sequence 217248,
26	49	92.5	152	16	US-10-425-115-217250	Sequence 217250,
27	49	92.5	152	16	US-10-425-115-217255	Sequence 217255,
28	49	92.5	152	16	US-10-425-115-218998	Sequence 218998,
29	49	92.5	152	16	US-10-425-115-218999	Sequence 218999,
30	49	92.5	152	16	US-10-425-115-219001	Sequence 219001,
31	49	92.5	152	16	US-10-425-115-219004	Sequence 219004,
32	49	92.5	152	16	US-10-425-115-219005	Sequence 219005,
33	49	92.5	152	16	US-10-425-115-219006	Sequence 219006,
34	49	92.5	152	16	US-10-425-115-289038	Sequence 289038,
35	49	92.5	152	16	US-10-425-115-289039	Sequence 289039,
36	49	92.5	152	16	US-10-425-115-289047	Sequence 289047,
37	49	92.5	152	17	US-10-732-923-18523	Sequence 18523, A
38	49	92.5	157	15	US-10-425-114-57694	Sequence 57694, A
39	49	92.5	162	16	US-10-425-115-217251	Sequence 217251,
40	49	92.5	166	15	US-10-264-049-3515	Sequence 3515, Ap
41	49	92.5	171	15	US-10-425-114-44835	Sequence 44835, A
42	49	92.5	172	15	US-10-264-049-3306	Sequence 3306, Ap
43	49	92.5	176	16	US-10-437-963-148926	Sequence 148926,
44	49	92.5	181	16	US-10-425-115-289043	Sequence 289043,
45	49	92.5	191	15	US-10-424-599-158043	Sequence 158043,

ALIGNMENTS

RESULT 1

US-10-006-177-2

Sequence 2, Application US/10006177

Publication No. US20030165513A1

GENERAL INFORMATION:

APPLICANT: Ramakrishna, Venky

APPLICANT: Philip, Ramila

TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treatm

TITLE OF INVENTION: Diagnosis of Cancer

FILE REFERENCE: 26747-35

CURRENT APPLICATION NUMBER: US/10/006.177

CURRENT FILING DATE: 2001-12-04

PRIOR APPLICATION NUMBER: US/60/251,022

PRIOR FILING DATE: 2000-12-04

PRIOR APPLICATION NUMBER: US/60/256,824

PRIOR FILING DATE: 2000-12-20

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn version 3.0

SEQ ID NO 2

LENGTH: 10

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Epitopic Peptide

US-10-006-177-2

Query Match 100.0%; Score 53; DB 14; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.014;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10

Db 1 RLQEDPPAGV 10

```
RESULT 2
US-10-425-115-234519
; Sequence 234519, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 234519
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Zea mays
; NAME/KEY: unsure
; LOCATION: (1)..(89)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_145467C.1.pep
US-10-425-115-234519

Query Match 100.0%; Score 53; DB 16; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLOEDPPAGV 10
Db 2 RLOEDPPAGV 11

RESULT 3
US-10-489-695-15
; Sequence 15, Application US/10489695
; Publication No. US20050107293A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; SPRAGUE, William W.;
; APPLICANT: CHAWLA, Narinder K.; WARREN, Bridget A.;
; APPLICANT: TANG, Y. Tom; ELLIOTT, Vicki S.;
; APPLICANT: MARQUIS, Joseph P.; LI, Joana X.;
; APPLICANT: GRIFFIN, Jennifer A.; GIETZEN, Kimberly J.;
; APPLICANT: YANG, Junming; LU, Dyung Aina M.;
; APPLICANT: EMERLING, Brooke M.; DUGGAN, Brendan M.;
; APPLICANT: RICHARDSON, Thomas W.; LEE, Soo Yeun;
; APPLICANT: RAMKUMAR, Jayalaxmi; BECHA, Shanya D.;
; APPLICANT: LEHR-MASON, Patricia M.; SWARNAKAR, Anita;
; APPLICANT: TRAN, Uyen K.; KABLE, Amy E.;
; APPLICANT: HAFALIA, April J.A.; KHARE, Reena
; TITLE OF INVENTION: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES
; FILE REFERENCE: PF-1186 USN
; CURRENT APPLICATION NUMBER: US/10/489,695
; CURRENT FILING DATE: 2004-03-15
; PRIOR APPLICATION NUMBER: PCT/US02/29221
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/322,196
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/324,134
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/327,233
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/346,198
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/343,980
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 60/348,887
; PRIOR FILING DATE: 2001-11-09

US-10-425-115-234519
; PRIOR APPLICATION NUMBER: US 60/332,423
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/334,145
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/334,229
; PRIOR FILING DATE: 2001-11-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No: 7503485CDI
US-10-489-695-15

Query Match 100.0%; Score 53; DB 17; Length 122;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLOEDPPAGV 10
Db 15 RLOEDPPAGV 24

RESULT 4
US-10-732-923-18529
; Sequence 18529, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 18529
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Equus caballus
US-10-732-923-18529

Query Match 100.0%; Score 53; DB 17; Length 126;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLOEDPPAGV 10
Db 6 RLOEDPPAGV 15

RESULT 5
US-10-732-923-18530
; Sequence 18530, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 18530
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-732-923-18530
```



Query Match 100.0%; Score 53; DB 17; Length 126;  
Best Local Similarity 100.0%; Pred. No. 0.19;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 6 RLQEDPPAGV 15

RESULT 6  
US-10-116-275-164  
; Sequence 164, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116,275  
; CURRENT FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 164  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-275-164

Query Match 100.0%; Score 53; DB 15; Length 152;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPAGV 24

RESULT 7  
US-10-425-115-356888  
; Sequence 356888, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 356888  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_88650C.1.pap  
US-10-425-115-356888

Query Match 92.5%; Score 49; DB 16; Length 58;  
Best Local Similarity 80.0%; Pred. No. 0.43;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPAGV 24

RESULT 8  
US-10-425-115-217253  
; Sequence 217253, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 217253  
; LENGTH: 69  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(69)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_129736C.1.pap  
US-10-425-115-217253

Query Match 92.5%; Score 49; DB 16; Length 69;  
Best Local Similarity 80.0%; Pred. No. 0.52;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPAGV 24

RESULT 9  
US-10-424-599-145487  
; Sequence 145487, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 145487  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(73)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_102394C.1.pap  
US-10-424-599-145487

Query Match 92.5%; Score 49; DB 15; Length 73;  
Best Local Similarity 80.0%; Pred. No. 0.55;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQEDPPAGV 24

RESULT 10  
US-10-424-599-282590  
; Sequence 282590, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 282590  
; LENGTH: 77  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(77)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_971C.1.pep  
US-10-424-599-282590

Query Match 92.5%; Score 49; DB 15; Length 77;  
Best Local Similarity 80.0%; Pred. No. 0.58;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
| | | : | | | | |  
Db 14 RLQDPPAGI 23

RESULT 11  
US-10-424-599-194713  
; Sequence 194713, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 194713  
; LENGTH: 82  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(82)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_17853C.1.pep  
US-10-424-599-194713

Query Match 92.5%; Score 49; DB 15; Length 82;  
Best Local Similarity 80.0%; Pred. No. 0.62;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
| | | : | | | | |  
Db 15 RLQDPPAGI 24

RESULT 12  
US-10-424-599-169060  
; Sequence 169060, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 169060  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(83)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_123676C.1.pep  
US-10-424-599-169060

Query Match 92.5%; Score 49; DB 15; Length 83;  
Best Local Similarity 80.0%; Pred. No. 0.63;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
| | | : | | | | |  
Db 12 RLQDPPAGI 21

RESULT 13  
US-10-425-115-218996  
; Sequence 218996, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 218996  
; LENGTH: 93  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(93)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_131316C.1.pep  
US-10-425-115-218996

Query Match 92.5%; Score 49; DB 16; Length 93;  
Best Local Similarity 80.0%; Pred. No. 0.7;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
| | | : | | | | |  
Db 15 RLQDPPAGI 24

RESULT 14

US-10-425-115-225960  
; Sequence 225960, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 225960  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_137666C.1.pep  
US-10-425-115-225960

Query Match 92.5%; Score 49; DB 16; Length 115;  
Best Local Similarity 80.0%; Pred. No. 0.88;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQDPPAGI 24

RESULT 15  
US-10-425-115-217252  
; Sequence 217252, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 217252  
; LENGTH: 121  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_129735C.1.pep  
US-10-425-115-217252

Query Match 92.5%; Score 49; DB 16; Length 121;  
Best Local Similarity 80.0%; Pred. No. 0.93;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLQEDPPAGV 10  
Db 15 RLQDPPAGI 24

Search completed: November 18, 2005, 10:44:29  
Job time : 51.2473 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:03:06 ; Search time 48.7742 Seconds  
(without alignments)  
71.366 Million cell updates/sec

Title: US-10-006-177-3

Perfect score: 44

Sequence: 1 KLDVGNARV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	9	5	ABG32303 HLA-A2 as
2	44	100.0	87	7	ADJ70645 Human hea
3	44	100.0	96	5	ABD06075 Human NS
4	44	100.0	123	4	ABG28199 Novel hum
5	44	100.0	159	8	ADP04246 Human col
6	44	100.0	246	2	AAW77312 Polypepti
7	44	100.0	246	8	ADP05470 Human mem
8	44	100.0	246	8	ABM82472 Tumour-as
9	44	100.0	246	8	ADP23037 PRO polyp
10	44	100.0	291	5	AAU69428 Lung smal
11	44	100.0	301	8	ADP04248 Human col
12	44	100.0	313	6	ADA55664 Human pro
13	44	100.0	313	8	ADP04247 Human col
14	44	100.0	313	8	ADQ66364 Novel hum
15	37	84.1	282	6	ABP79861 N. gonorr
16	36	81.8	245	7	ADB85218 Mouse BAP
17	34	77.3	275	6	ABU38231 Protein e
18	34	77.3	352	7	ABO71243 Pseudomon
19	34	77.3	435	7	ABO79761 Pseudomon
20	33	75.0	460	4	ABB62929 Pseudophil
21	33	75.0	783	7	ABO83357 Pseudomon
22	33	75.0	839	4	ABB63206 Drosophil
23	32	72.7	168	2	AAW06596 Human end
24	32	72.7	168	3	AAV69674 Human end
25	32	72.7	171	6	ABU43429 Protein e

26	32	72.7	178	4	AAAB47613 Human Tyr
27	32	72.7	178	5	AAEI3489 Human Tyr
28	32	72.7	178	6	ABU72382 Human Tyr
29	32	72.7	196	1	AAAP70441 Sequence
30	32	72.7	196	1	AAAP82979 Part of c
31	32	72.7	216	3	AAAB58144 Lung canc
32	32	72.7	216	3	AAAB58144 Lung canc
33	32	72.7	240	8	ADN99944 Novel hum
34	32	72.7	275	7	ADH86428 Enterococ
35	32	72.7	301	2	AAW53962 Monocyte
36	32	72.7	301	8	ADE76967 Human pro
37	32	72.7	388	6	ABU39484 Protein e
38	32	72.7	456	8	ADI67184 Lactobaci
39	32	72.7	494	2	AAW47017 Arabidopsi
40	32	72.7	527	7	ABU64299 Human tyr
41	32	72.7	528	7	ADF76697 Novel hum
42	32	72.7	528	8	ADN05830 Antipsori
43	32	72.7	528	8	ADQ65849 Novel hum
44	32	72.7	528	8	ABM80058 Tumour-as
45	32	72.7	528	8	ADP24469 PRO polyp
			534	8	ADQ43575 Bacterial

## ALIGNMENTS

RESULT 1  
ABG32303  
ID ABG32303 standard; peptide; 9 AA.

XX AC ABG32303;

XX DT 05-NOV-2002 (first entry)

XX DE HLA-A2 associated immunogenic peptide from human BAP31 protein.

XX KW Human; immunogen; epitope: HLA-A1; human leukocyte antigen; CTL;  
XX KW cytotoxic Y lymphocyte; cytostatic; cancer; colorectal carcinoma;  
XX KW ovarian carcinoma; lung carcinoma; prostate carcinoma; vaccine; tumour;  
XX KW HLA-2; passive immunotherapy; BAP31; accessory protein.

XX OS Homo sapiens.

XX PN WO200246416-A2.

XX PD 13-JUN-2002.

XX PF 04-DEC-2001; 2001WO-US047290.

XX PR 04-DEC-2000; 2000US-0251022P.

XX XX 20-DEC-2000; 2000US-0256824P.

XX FA (ARGO-) ARGONEX INC.

XX PI Ramakrishna V, Ross M, Philip R;

XX XX WPI; 2002-619021/66.

XX XX New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte,  
XX XX and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.

XX XX Claim 1; Page 50; 60pp; English.

XX CC The invention relates to an immunogen comprising an isolated polypeptide  
XX CC whose amino acid sequence comprises an epitopic peptide, does not include  
XX CC MAGE 4 or MFG-E8 proteins, or consists of MAGE D protein or its  
XX CC immunologically active fragment. Also included are a polynucleotide  
XX CC encoding the immunogen or its complement, a vector comprising the  
XX CC polynucleotide, a mammalian cell comprising the vector and expressing the  
XX CC polynucleotide, a vaccine composition comprising the immunogen and an  
XX CC antibody specific for the immunogen. The immunogen is useful for inducing  
XX CC a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour  
XX CC cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is  
XX CC useful for inducing a CTL response when administered to a subject. A

CC mammalian cell that can express the immunogen, is useful for inducing a  
 CC CTL response in vitro that is specific for a tumour cell expressing HLA-1  
 CC or HLA-2. The immunogen or cell is useful for inducing CTL for treating a  
 CC subject with cancer (carcinoma, preferably colorectal carcinoma, ovarian  
 CC carcinoma, lung carcinoma and prostate carcinoma). The immunogen is also  
 CC useful for screening and diagnostic agents, for gene screening in  
 CC patients afflicted with cancer, for screening a sample for the presence  
 CC of CTUs that specifically recognise the corresponding epitopes, as a  
 CC diagnostic tool to evaluate the efficacy of the immunotherapeutic  
 CC treatments, to prepare class I MHC (major histocompatibility class)  
 CC tetramers which are utilised in conjunction with flow cytometry to  
 CC quantitate the frequency of peptide-specific CTL that are present in a  
 CC population of lymphocytes from an individual, and for stimulating the  
 CC production of antibodies for use in passive immunotherapy, for use as  
 CC diagnostic reagents, and for use as reagents in other processes such as  
 CC affinity chromatography. The present sequence is an immunogenic epitope  
 CC of the invention derived from human BAP1 (accessory protein)

SQ Sequence 9 AA;

Query Match 100.0%; Score 44; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
 |||||  
 Db 1 KLDVGNAEV 9

RESULT 2

ADJ70645

ID ADJ70645 standard; protein; 87 AA.

XX AC ADJ70645;

XX DT 06-MAY-2004 (first entry)

XX DE Human heat mitochondrial protein as a therapeutic target SeqID2451.

XX KW mitochondrial; human; screening assay; diabetes mellitus;  
 KW Huntington's disease; osteoarthritis;  
 KW Leber's hereditary optic neuropathy; LHON;  
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
 KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;  
 KW osteopathic; ophthalmological; cytostatic.

XX OS Homo sapiens.

XX FN WO2003087768-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010870.

XX PR 12-APR-2002; 2002US-0372843P.

XX PR 17-JUN-2002; 2002US-0389987P.

XX PR 20-SEP-2002; 2002US-0412418P.

XX PA (MITO-) MITOKOR.

XX PA (BUCK-) BUCK INST AGE RES.

XX PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;

XX PI Warnock DE;

XX DR WPI; 2003-845369/78.

XX XX Identifying a mitochondrial target for drug screening assays and for

PT treating diseases associated with altered mitochondrial function,

PT comprises detecting a modified polypeptide in a sample and correlating

PT with the disease.

XX Claim 1; SEQ ID NO 2451; 180pp; English.

XX CC This invention relates to novel mitochondrial targets that can be used  
 CC for therapeutic intervention in treating a disease associated with  
 CC altered mitochondrial function. Specifically, it refers to a method for  
 CC identifying proteins of the human heart mitochondrial proteome that are  
 CC useful for drug screening assays, as well as therapeutic targets. The  
 CC present invention describes a method for identifying such proteins that  
 CC can be used in the treatment of various diseases associated with altered  
 CC mitochondrial function including diabetes mellitus, Huntington's disease,  
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 CC compositions have neuroprotective, nontropic, antidiabetic,  
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 CC cyostatic activities. This polypeptide sequence is a human heart  
 CC mitochondrial protein of the invention.

SQ Sequence 87 AA;

Query Match 100.0%; Score 44; DB 7; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 0.14; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;

Qy 1 KLDVGNAEV 9  
 |||||  
 Db 27 KLDVGNAEV 35

RESULT 3

ABB06075

ID ABB06075 standard; protein; 96 AA.

XX AC ABB06075;

XX DT 10-MAY-2002 (first entry)

XX DE Human NS protein sequence SEQ ID NO:167.

XX KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;  
 KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;  
 KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;  
 KW anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;  
 KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiac;  
 KW anticonvulsant; antidiabetic; tranquiliser; antidepressant; neuroleptic;  
 KW gastrointestinal; virucide; antiulcer; cerebroprotective; nontropic;  
 KW contraceptive; vaccine; Gene therapy; cancer; osteoporosis; dystonia;  
 KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;  
 KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;  
 KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;  
 KW infertility; cardiovascular disease; coagulation disease; hypertension;  
 KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;  
 KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;  
 KW gastric ulcer; Alzheimer's disease.

XX OS Homo sapiens.

XX OS WO200206315-A2.

XX PD 24-JAN-2002.

XX PF 17-JUL-2001; 2001WO-IL000653.

XX PR 18-JUL-2000; 2000IL-00137345.

XX PR 15-DEC-2000; 2000IL-00140354.

XX PA (COMP-) COMPUGEN LTD.

XX PI Mintz L, Freilich S, Bernstein J;

XX DR WPI; 2002-155037/20.

XX DR N-PSDB; ABL39729.

PT One hundred and twenty eight novel nucleic acid sequences, useful for

XX treating and diagnosing e.g. cancer, asthma and Alzheimer's.  
 PS Claim 6; Page 187; 290pp; English.  
 XX ABL39691 to ABL39818 represent novel human nucleic acid sequences  
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences  
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,  
 CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,  
 CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,  
 CC anorectic, muscular, anti-HIV, antifertility, cardiovascular,  
 CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,  
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,  
 CC antidepressant, gastrointestinal, neurolaptic, cerebroprotective,  
 CC neotropic and contraceptive activities. The NS can be used in vaccines,  
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and  
 CC antibodies from the present invention can be used for treating and  
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative  
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,  
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,  
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular  
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune  
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,  
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,  
 CC Alzheimer's disease and as a contraceptive  
 XX

XX Sequence 96 AA;

Query Match 100.0%; Score 44; DB 5; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
 Db 17 KLDVGNAEV 25  
 |||||

RESULT 4

ABG28199  
 ID ABG28199 standard; protein; 123 AA.

XX AC ABG28199;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #28190.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS92386.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

PS Claim 20; SEQ ID NO 58558; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX

XX Sequence 123 AA;

Query Match 100.0%; Score 44; DB 4; Length 123;  
 Best Local Similarity 100.0%; Pred. No. 0.21;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
 Db 43 KLDVGNAEV 51  
 |||||

RESULT 5

ADP04246  
 ID ADP04246 standard; protein; 159 AA.

XX AC ADP04246;

XX 09-SEP-2004 (first entry)

XX Human colon specific protein SEQ ID NO:227.

XX human; colon specific nucleic acid; CSNA; colon specific protein; CSP;  
 KW cytostatic; vaccine; gene therapy; colon cancer.

XX Homo sapiens.

XX WO2004050858-A2.

XX 17-JUN-2004.

XX 04-DEC-2003; 2003WO-US038808.

XX 04-DEC-2002; 2002US-0431133P.

XX (DIAD-) DIADEXUS INC.

XX Macina RA, Turner LR, Sun Y, Rodriguez M, Burcham TS;

XX WPI; 2004-480622/45.

XX Novel colon specific protein derived from normal and neoplastic colon  
 PT cell, useful as vaccine in treating colon cancer and in identifying,  
 PT diagnosing, monitoring, staging, imaging colon cancer and non-cancerous  
 PT disease state in colon.

XX Claim 12; SEQ ID NO 227; 655pp; English.

XX The invention relates to a novel colon specific protein (CSP) (I), and  
 CC the nucleic acid encoding it. A CSP of the invention has cytostatic  
 CC activity, and may have use in a vaccine, and in gene therapy. The CSP is

CC useful for determining the presence of a colon specific protein in a  
 CC sample. The nucleic acid encoding the CSP is useful for determining the  
 CC presence of a colon specific nucleic acid (CSNA) in a sample. The CSP and  
 CC CSNA are useful for diagnosing or monitoring the presence and metastases  
 CC of colon cancer in a patient. The method of administering a composition  
 CC comprising a CSP or CSNA is useful for treating a patient with colon  
 CC cancer. The CSP and CSNA are useful as vaccine for treating colon cancer  
 CC and non-cancerous disease states in colon. The present sequence  
 CC represents a CSP of the invention.

XX Sequence 159 AA;

Query Match 100.0%; Score 44; DB 8; Length 159;  
 Best Local Similarity 100.0%; Pred. NO. 0.28;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
 Db |||||  
 80 KLDVGNAEV 88

RESULT 6  
 AAW77312  
 ID AAW77312 standard; protein; 246 AA.

XX AC AAW77312;

DT 16-DEC-1998 (first entry)

XX Polypeptide sequence of p28 Bap31/CDM.

XX Apoptosis; p28 Bap31/CDM; modulate; p28 Bap31; human; p20 product;  
 KW pro-FLICE/Bcl-2 protein; Bcl-2; Bcl-XL; degenerative disease; neoplasia.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 2..11  
 FT /note= "sequence homologous to human Bap31"  
 FT Domain 7..27  
 FT /note= "predicted transmembrane domain TM1"  
 FT Domain 43..64  
 FT /note= "predicted transmembrane domain TM2"  
 FT Domain 102..123  
 FT /note= "predicted transmembrane domain TM3"  
 FT Peptide 161..165  
 FT /note= "caspase recognition site"  
 FT Cleavage-site 164..165  
 FT Peptide 235..239  
 FT /note= "caspase recognition site"  
 FT Cleavage-site 238..239  
 FT Peptide 243..246  
 FT /note= "ER retention signal sequence"

XX WO9839434-A1.

XX 11-SEP-1998.

XX 02-MAR-1998; 98WO-1B000706.

XX 03-MAR-1997; 97CA-02198988.

XX (UYWC-) UNIV MCGILL.

XX Shore GC, Ng WHF, Nguyen M, Branton PE;

XX WPI; 1998-506359/43.

XX Polypeptide modulating apoptosis, p28 Bap31 - useful to diagnose diseases  
 PT involving altered apoptosis, e.g. degenerative diseases or neoplasia and  
 PT to detect compounds modulating apoptosis.

XX Claim 7; Fig 2C; 130pp; English.

XX This represents the amino acid sequence of the p28 Bap31/CDM polypeptide  
 CC that can modulate apoptosis. The polypeptide p28 Bap31, fragments and  
 CC encoding nucleic acids are useful to generate or identify compounds which  
 CC alter the biological activity of p28 Bap31 and thus modulate apoptosis  
 CC when administered to cells (e.g. mammalian and especially human or rodent  
 CC cells). Such compounds may affect e.g. p28 Bap31 cleavage to produce p20  
 CC product, formation of complex with pro-FLICE/Bcl-2 protein (especially  
 CC Bcl-2 or Bcl-XL), specific binding to an antibody or p28 Bap31 expression  
 CC in the cell. The compounds are especially useful to inhibit apoptosis in  
 CC mammals with degenerative diseases or to increase apoptosis in mammals  
 CC with neoplasia. p28 Bap31 can also be used diagnostically, by measuring  
 CC levels in mammalian samples whereby a reduction is indicative of the  
 CC presence or likelihood of a disease caused by decreased apoptosis  
 CC (neoplasia) and an increase indicates the presence of a disease caused by  
 CC increased apoptosis (especially a degenerative disease). Antibodies  
 CC specifically binding p28 Bap31/p20 can similarly be used to diagnose  
 CC diseases/likelihood of diseases involving altered apoptosis

XX Sequence 246 AA;

Query Match 100.0%; Score 44; DB 2; Length 246;  
 Best Local Similarity 100.0%; Pred. NO. 0.47;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
 Db |||||  
 167 KLDVGNAEV 175

RESULT 7

ADP05470  
 ID ADP05470 standard; protein; 246 AA.

XX AC ADP05470;

XX 29-JUL-2004 (first entry)

XX Human membrane IgD binding protein 31.

XX anorectic; antidiabetic; antilipemic; hypotensive; antiarteriosclerotic;  
 KW vasotropic; disease marker; neutral lipid; neutral lipid accumulation;  
 KW obesity; fatty liver; type II diabetes; hyperlipidaemia; hypertension;  
 KW arteriosclerosis; ischaemic heart disease; human;  
 KW membrane IgD binding protein 31.

XX Homo sapiens.

XX JP2004129587-A.

XX 30-APR-2004.

XX 11-OCT-2002; 2002JP-00298273.

XX 11-OCT-2002; 2002JP-00298273.

XX (SUMU ) SUMITOMO SEIYAKU KK.

XX WPI; 2004-395156/37.

XX N-PSDB; ADP05451.

XX New polynucleotide, useful as disease marker of disease accompanying  
 PT accumulation of neutral lipid such as obesity and hypertension.

XX Claim 4; SEQ ID NO 36; 146pp; Japanese.

XX The invention describes a disease marker of disease accompanying  
 CC accumulation of neutral lipid, comprising a polynucleotide having 15  
 CC continuous bases of any one of 19 fully defined sequence (S1) as given in  
 CC the specification, and/or a polynucleotide which is complementary to the  
 CC above mentioned polynucleotide, is new. (I) is useful as a probe or  
 CC primer for detecting disease accompanying accumulation of neutral lipid.  
 CC The method of detection involves preparing RNA from the biological sample



CC of a subject, reacting the complementary polynucleotide and (I),  
 CC measuring the disease marker for the complementary polynucleotide  
 CC transferred from RNA derived from biological sample specifically coupled  
 CC with the disease marker, and detecting the disease accompanying storage  
 CC of neutral lipid by preparing an index for the amount of coupling of  
 CC disease marker as compared with the result of measurement in a normal  
 CC biological sample. The diseases include obesity, fatty liver, type II  
 CC diabetes, hyperlipidemia, hypertension, arteriosclerosis and ischaemic  
 CC heart disease. This is the amino acid sequence of human membrane IgB  
 CC binding protein 31, a protein associated with disease accompanying  
 CC accumulation of neutral lipid.  
 XX  
 SQ Sequence 246 AA;

Query Match 100.0%; Score 44; DB 8; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 0.47;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLDVGNAEV 9  
 |||||  
 Db 167 KLDVGNAEV 175

## RESULT 8

ID ABM82472 standard; protein; 246 AA.  
 XX AC ABM82472;

DT 18-NOV-2004 (first entry)

DE Tumour-associated antigenic target (TAT) polypeptide PRO4885, SEQ:6345.

KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;  
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
 KW central nervous system cancer; bladder cancer; pancreatic cancer;  
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;  
 KW chromosome identification; chromosome mapping; gene mapping;  
 KW gene therapy; cytostatic.

OS Homo sapiens.

PN WO2004030615-A2.

PD 15-APR-2004.

PF 29-SEP-2003; 2003WO-US028547.

PR 02-OCT-2002; 2002US-0414971P.

PA (GETH ) GENENTECH INC.

PI Wu TD, Zhang Z, Zhou Y;

DR WPI; 2004-347921/32.

DR N-PSDB; ACN41116.

PT New tumor-associated antigenic target polypeptides and nucleic acids,  
 PT useful in preparing a medicament for treating or detecting a  
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
 PT prostate cancer or tumor.

XX Claim 12; SEQ ID NO 6345; 7273pp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)  
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
 CC overexpressed in cancer tissues compared to normal tissues, and may thus  
 CC serve as effective targets for the diagnosis and treatment of cancer in  
 CC mammals. The invention also relates to nucleic acid and polypeptide  
 CC sequences at least 80% identical to the TAT nucleic acids and  
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic

CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
 CC TAT polypeptide; and methods and compositions for the treatment or  
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
 CC antibodies, antagonists, binding molecules and compositions are useful  
 CC for diagnosing or treating a cell proliferative disorder associated with  
 CC increased TAT expression, particularly cancers such as breast cancer,  
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
 CC used as hybridisation probes, in chromosome and gene mapping, in  
 CC chromosome identification and in gene therapy. The present sequence  
 XX represents a TAT polypeptide of the invention

SQ Sequence 246 AA;

Query Match 100.0%; Score 44; DB 8; Length 246;

Best Local Similarity 100.0%; Pred. No. 0.47;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLDVGNAEV 9

|||||

Db 167 KLDVGNAEV 175

## RESULT 9

ID ADP23037 standard; protein; 246 AA.  
 XX AC ADP23037;

DT 18-NOV-2004 (first entry)

DE PRO polypeptide SEQ ID NO:131.

KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;  
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;  
 KW antiasmatic; hepatotropic; respiratory; gene therapy; immune system.

OS Unidentified.

PN WO2004041170-A2.

PD 21-MAY-2004.

PF 30-OCT-2003; 2003WO-US034312.

PR 01-NOV-2002; 2002US-0423394P.

PA (GETH ) GENENTECH INC.

PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
 PI Wu TD;

DR WPI; 2004-419628/39.

DR N-PSDB; ADP23036.

PT New PRO polypeptides and polynucleotides, useful for treating e.g.  
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
 PT renal disease, or demyelinating diseases of the central or peripheral  
 PT nervous system.

XX Claim 7; SEQ ID NO 131; 2940pp; English.

XX The invention relates to a novel isolated nucleic acid and the PRO  
 CC polypeptide encoded by it. A protein of the invention has

CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
 CC antiasmatic, hepatotropic, and respiratory activity. A polynucleotide  
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
 CC agonist, antagonist, or antibody that specifically binds to the  
 CC polypeptide is useful for treating an immune related disorder such as  
 CC systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,  
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an

CC idiopathic inflammatory myopathy. Sjogren's syndrome, systemic  
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
 CC disease, a demyelinating disease of the central or peripheral nervous  
 CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
 CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
 CC disease, infectious or autoimmune chronic active hepatitis, primary  
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
 CC hypersensitivity, urticaria, an immunologic disease of the lung,  
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
 CC pneumonitis, a transplantation associated disease, graft rejection or  
 CC graft-versus-host disease. The present sequence represents a PRO protein  
 CC of the invention.

XX  
 XX  
 SQ Sequence 246 AA;

Query Match 100.0%; Score 44; DB 8; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 0.47;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
 |||||  
 Db 167 KLDVGNAEV 175

RESULT 10  
 AAU69428  
 ID AAU69428 standard; protein; 291 AA.

XX AC AAU69428;

XX 30-JAN-2002 (first entry)

XX Lung small cell carcinoma antigen #22.

XX Human; cytostatic; antitumour; lung small cell cancer antigen; tumour;  
 XX lung cancer.

XX Homo sapiens.

XX WO200177168-A2.

XX 18-OCT-2001.

XX 11-APR-2001; 2001WO-US011859.

XX 11-APR-2000; 2000US-0196780P.

XX 21-JUN-2000; 2000US-0213361P.

XX 01-SEP-2000; 2000US-0229763P.

XX 05-SEP-2000; 2000US-0230629P.

XX 14-SEP-2000; 2000US-0232565P.

XX 19-DEC-2000; 2000US-0257037P.

XX 08-JAN-2001; 2001US-0260796P.

XX (CORI-) CORIXA CORP.

XX Lodes MJ, Wang T, Mohamath R, Indirias CY;

XX WPI; 2002-010896/01.

XX N-PSDB; AAS61868.

XX Lung tumor polynucleotide and polypeptides useful in therapy and  
 XX diagnosis of cancer especially lung cancer.

XX Claim 2; Page 290-291; 295pp; English.

XX The invention relates to novel isolated lung small cell cancer antigen  
 CC polynucleotides (I) and polypeptides (II) used in a method of detecting  
 CC cancer in a patient. The method is optionally performed by utilising

CC oligonucleotides (III), where the biological sample from the patient is  
 CC contacted with (III), detecting the amount of polynucleotide hybridised  
 CC to (III) in the sample and comparing the amount of polynucleotide to a  
 CC predetermined cut-off value and thereby determining cancer in a patient.  
 CC (I), (II) or antigen-presenting cells expressing (II) is useful for  
 CC stimulating and/or expanding T cells specific for a tumour protein. The  
 CC method comprises contacting T cells with one of the components under  
 CC conditions to permit the stimulation and/or expansion of the cells. A  
 CC composition comprising (I) is useful for stimulating an immune response  
 CC in a patient and for inhibiting the development of a cancer especially  
 CC lung cancer in a patient. An isolated T cell population is useful for  
 CC removing tumour cells from the biological sample and for inhibiting the  
 CC development of cancer in a patient. AAU69407-AAU69431 represent novel  
 CC human lung small cell cancer antigen amino acid sequences of the  
 CC invention

XX  
 XX  
 SQ Sequence 291 AA;

Query Match 100.0%; Score 44; DB 5; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 0.57;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
 |||||  
 Db 212 KLDVGNAEV 220

RESULT 11  
 ADP04248

ID ADP04248 standard; protein; 301 AA.

XX AC ADP04248;

XX 09-SEP-2004 (first entry)

XX Human colon specific protein SEQ ID NO:229.

XX human; colon specific nucleic acid; CSNA; colon specific protein; CSP;  
 XX cytostatic; vaccine; gene therapy; colon cancer.

XX Homo sapiens.

XX WO2004050858-A2.

XX 17-JUN-2004.

XX 04-DEC-2003; 2003WO-US038808.

XX 04-DEC-2002; 2002US-0431133P.

XX (DIAD-) DIADEXUS INC.

XX Macina RA, Turner LR, Sun Y, Rodriguez M, Burcham TS;

XX WPI; 2004-480622/45.

XX Novel colon specific protein derived from normal and neoplastic colon  
 cell, useful as vaccine in treating colon cancer and in identifying,  
 PT diagnosing, monitoring, staging, imaging colon cancer and non-cancerous  
 PT disease state in colon.

XX Claim 12; SEQ ID NO 229; 655pp; English.

XX The invention relates to a novel colon specific protein (CSP) (I), and  
 CC the nucleic acid encoding it. A CSP of the invention has cytostatic  
 CC activity, and may have use in a vaccine, and in gene therapy. The CSP is  
 CC useful for determining the presence of a colon specific protein in a  
 CC sample. The nucleic acid encoding the CSP is useful for determining the  
 CC presence of a colon specific nucleic acid (CSNA) in a sample. The CSP and  
 CC CSNA are useful for diagnosing or monitoring the presence and metastases  
 CC of colon cancer in a patient. The method of administering a composition  
 CC comprising a CSP or CSNA is useful for treating a patient with colon  
 CC cancer. The CSP and CSNA are useful as vaccine for treating colon cancer

CC and non-cancerous disease states in colon. The present sequence  
 CC represents a CSP of the invention.

XX Sequence 301 AA;

Query Match 100.0%; Score 44; DB 8; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 0.59; 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;

Qy 1 KLDVGNAEV 9

Db 167 KLDVGNAEV 175

RESULT 12

ADAS5664  
 ID ADAS5664 standard; protein; 313 AA.

XX AC ADAS5664;

XX DT 20-NOV-2003 (first entry)

XX DE Human protein, SEQ ID 3232.

XX KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
 KW inflammatory disease; osteoporosis; neurological disease.

XX OS Homo sapiens.

XX PN EP1293569-A2.

XX PD 19-MAR-2003.

XX PF 21-MAR-2002; 2002EP-00006586.

XX PR 14-SEP-2001; 2001JP-00328381.

XX PR 24-JAN-2002; 2002US-0350435P.

XX PA (HELI-) HELIX RES INST.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

XX PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX DR WPI; 2003-395539/38.

XX DR N-PSDB; ADAS4025.

XX PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.

XX PS Claim 14; SEQ ID NO 323; 205pp; English.

XX CC The present invention relates to novel human secretory or membrane  
 CC proteins (ADAS4072-ADAS5710) and their coding sequences (ADAS2433-  
 CC ADAS4071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 313 AA;

Query Match 100.0%; Score 44; DB 6; Length 313;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9

Db 234 KLDVGNAEV 242

RESULT 13

osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
 gene therapy; diagnostic marker; morbid state; osteoporosis;  
 neurological disease; Alzheimer's disease; Parkinson's disease; dementia;  
 cancer.

ADP04247

ID ADP04247 standard; protein; 313 AA.

XX AC ADP04247;

XX DT 09-SEP-2004 (first entry)

XX DE Human colon specific protein SEQ ID NO:228.

XX KW human; colon specific nucleic acid; CSNA; colon specific protein; CSP;  
 KW cytostatic; vaccine; gene therapy; colon cancer.

XX OS Homo sapiens.

XX PN WO2004050858-A2.

XX PD 17-JUN-2004.

XX PF 04-DEC-2003; 2003WO-US038808.

XX PR 04-DEC-2002; 2002US-0431133P.

XX PA (DIAD-) DIADEXUS INC.

XX PI Macina RA, Turner LR, Sun Y, Rodriguez M, Burcham TS;

XX DR WPI; 2004-480622/45.

XX PT Novel colon specific protein derived from normal and neoplastic colon  
 PT cell, useful as vaccine in treating colon cancer and in identifying,  
 PT diagnosing, monitoring, staging, imaging colon cancer and non-cancerous  
 PT disease state in colon.

XX PS Claim 12; SEQ ID NO 228; 655pp; English.

XX CC The invention relates to a novel colon specific protein (CSP) (I), and  
 CC the nucleic acid encoding it. A CSP of the invention has cytostatic  
 CC activity, and may have use in a vaccine, and in gene therapy. The CSP is  
 CC useful for determining the presence of a colon specific protein in a  
 CC sample. The nucleic acid encoding the CSP is useful for determining the  
 CC presence of a colon specific nucleic acid (CSNA) in a sample. The CSP and  
 CC CSNA are useful for diagnosing or monitoring the presence and metastases  
 CC of colon cancer in a patient. The method of administering a composition  
 CC comprising a CSP or CSNA is useful for treating a patient with colon  
 CC cancer. The CSP and CSNA are useful as vaccine for treating colon cancer  
 CC and non-cancerous disease states in colon. The present sequence  
 CC represents a CSP of the invention.

XX Sequence 313 AA;

Query Match 100.0%; Score 44; DB 8; Length 313;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9

Db 234 KLDVGNAEV 242

RESULT 14

ADQ66364  
 ID ADQ66364 standard; protein; 313 AA.

XX AC ADQ66364;

XX DT 07-OCT-2004 (first entry)

XX DE Novel human protein sequence #1337.

XX OS Homo sapiens.  
XX PN EPI440981-A2.  
XX PD 28-JUL-2004.  
XX PF 21-JAN-2004; 2004EP-00001196.  
XX PR 21-JAN-2003; 2003JP-00102206.  
XX PR 09-MAY-2003; 2003JP-00131392.  
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX PI Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Nagai K, Irie R;  
XX WPI: 2004-535376/52.  
DR N-PSDB; ADQ64176.  
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,  
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
XX Claim 1; SEQ ID NO 3525; 2449pp; English.  
XX The invention relates to 2495 novel polynucleotides (I) and their encoded  
CC polypeptides, sequences hybridizing to these nucleotides, sequences  
CC encoding partial polypeptides and sequences having 70% or 90% identity to  
CC the nucleotide and protein sequences. The nucleotides and polypeptides  
CC are useful as diagnostic markers or therapeutic target for the diseases  
CC or morbid states. They are also useful for treating osteoporosis,  
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,  
CC dementia and various cancers. This sequence corresponds to a protein  
CC sequence of the invention.  
XX SQ Sequence 313 AA;  
Query Match 100.0%; Score 44; DB 8; Length 313;  
Best Local Similarity 100.0%; Pred. No. 0.62;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KLDVGNAEV 9  
Db 234 KLDVGNAEV 242  
RESULT 15  
ID ABP79861 standard; protein; 282 AA.  
XX AC ABP79861;  
XX DT 07-MAR-2003 (first entry)  
XX DE N. gonorrhoeae amino acid sequence SEQ ID 6252.  
XX KW Antibacterial; infection; vaccine; gene therapy.  
XX OS Neisseria gonorrhoeae.  
XX PN WO200279243-A2.  
XX PD 10-OCT-2002.  
XX PF 12-FEB-2002; 2002WO-IB002069.  
XX PR 12-FEB-2001; 2001GB-00003424.  
XX PA (CHIR-) CHIRON SPA.  
XX PI Fontana MR, Pizza M, Masignani V, Monaci E;  
XX WPI: 2003-058415/05.  
DR N-PSDB; ABZ40831.  
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a  
PT medicament for treating or preventing N. gonorrhoeae infection.  
XX PS Disclosure; Page 643; 815pp; English.  
XX CC The present invention relates to proteins from Neisseria gonorrhoeae.  
CC Also disclosed are the nucleic acid molecules encoding the proteins and  
CC antibodies that specifically bind to the proteins. The composition  
CC comprising the protein, nucleic acid or antibody is useful for the  
CC manufacture of a medicament for treating or preventing N. gonorrhoeae  
CC infection, this may be in the form of a vaccine or gene therapy.  
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid  
CC molecules of the invention  
XX SQ Sequence 282 AA;  
Query Match 84.1%; Score 37; DB 6; Length 282;  
Best Local Similarity 66.7%; Pred. No. 19;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KLDVGNAEV 9  
Db 110 RLDIGNADV 118  
Search completed: November 18, 2005, 01:01:00  
Job time : 54.7742 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:11:06 ; Search time 8.90323 Seconds  
(without alignments)  
97.263 Million cell updates/sec

Title: US-10-006-177-3  
Perfect score: 44  
Sequence: 1 KLDVGNREV 9  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	246	2	CDM protein - huma
2	36	81.8	245	2	BCR-associated pro
3	35	79.5	530	2	probable high affi
4	35	79.5	1295	2	hypothetical prote
5	34	77.3	174	2	hypothetical prote
6	34	77.3	275	2	transcription regu
7	34	77.3	483	2	isocitrate dehydro
8	33	75.0	38	2	glyoxaldehyde-3-p
9	33	75.0	443	2	hypothetical prote
10	33	75.0	704	2	probable DNA-direc
11	32	72.7	65	2	hypothetical prote
12	32	72.7	331	2	pyridoxal phosphat
13	32	72.7	446	2	phosphopyruvate hy
14	32	72.7	483	2	isocitrate dehydro
15	32	72.7	525	2	hypothetical prote
16	32	72.7	534	2	probable membrane
17	32	72.7	556	2	hypothetical prote
18	32	72.7	601	2	hypothetical prote
19	32	72.7	1137	2	pyruvate carboxyla
20	31	70.5	148	2	hypothetical prote
21	31	70.5	198	2	protein T0604.5 li
22	31	70.5	198	2	hypothetical prote
23	31	70.5	334	2	replication protei
24	31	70.5	462	1	3-oxoacyl-[acyl-ca
25	31	70.5	482	2	NADP-dependent gly
26	31	70.5	507	2	LAT1 protein - hum
27	31	70.5	507	2	high affinity nitr
28	31	70.5	509	2	high affinity nitr
29	31	70.5	531	2	xpae protein - xan

xpae protein - xan  
probable chemotaxi  
high-affinity nitr  
acetolactate synth  
zinc finger protei  
xanthine dehydroge  
xanthine dehydroge  
hypothetical prote  
hypothetical prote  
hypothetical prote  
cnc regulatory pro  
cnc protein homolo  
hypothetical prote  
conserved hypothet  
hypothetical prote

## ALIGNMENTS

### RESULT 1

S44279  
CDM protein - human  
N:Alternate names: B-cell receptor-associated protein BAP31; tumor-associated antigen 6C  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text change 09-Jul-2004  
C:Accession: S44279; S68962; G02096; S71117; S47240; S49265  
R:Mosser, J.; Sarde, C.O.; Vicaire, S.; Yates, J.R.; Mandel, J.L.  
submitted to the EMBL Data Library, March 1994  
A:Description: A new human gene with ubiquitous expression, located in Xq28 adjacent to  
A:Reference number: S44279  
A:Accession: S44279  
A:Molecule type: mRNA  
A:Residues: 1-246 <MOS>  
A:Cross-references: UNIPROT:P51572; EMBL:Z31696; NID:G479156; PID:G479157  
A:Experimental source: fetal fibroblast  
R:Li, E.; Bestagno, M.; Burrone, O.  
Eur. J. Biochem. 238, 631-638, 1996  
A:Title: Molecular cloning and characterization of a transmembrane surface antigen in hu  
A:Reference number: S68962; MUID:96300225; PMID:8706661  
A:Accession: S68962  
A:Molecule type: mRNA  
A:Residues: 1-246 <LiE>  
A:Cross-references: EMBL:X81109; NID:G535057; PIDN:CAA57015.1; PID:G535058  
A:Experimental source: erythroleukemia cell line K-562  
R:Eichler, E.E.; Lu, F.; Muzny, D.; Gibbs, R.A.; Nelson, D.L.  
submitted to the EMBL Data Library, September 1995  
A:Reference number: G02096  
A:Accession: G02096  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 120-246 <EIC>  
A:Cross-references: EMBL:U36341; NID:G1020318; PID:G1020320  
R:Adachi, T.; Schamel, W.W.A.; Kim, K.M.; Watanabe, T.; Becker, B.; Nielsen, P.J.; Reth,  
EMBO J. 15, 1534-1541, 1996  
A:Title: The specificity of association of the IgD molecule with the accessory proteins  
A:Reference number: S71116; MUID:96203070; PMID:8612576  
A:Accession: S71117  
A:Molecule type: mRNA  
A:Residues: 1,T,3-207,E',209-246 <ADA>  
A:Cross-references: EMBL:X81817; NID:G550342; PIDN:CAA57415.1; PID:G550343  
A:Experimental source: pro-B cell line FLEB-14-14  
C:Genetics:  
A:Gene: CDM; BAP31  
A:Map position: Xq28  
A:Introns: 159/3; 201/1; 234/3  
A:Note: This list of introns may be incomplete  
C:Keywords: surface antigen; transmembrane protein  
F:5-27/Domain: (or 8-27) transmembrane #status predicted <TM1>  
F:51-67/Domain: transmembrane #status predicted <TM2>  
F:103-123/Domain: transmembrane #status predicted <TM3>

```

Query Match      100.0%; Score 44; DB 2; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 KLDVGNAEV 9
Db 167 KLDVGNAEV 175

RESULT 2
S71116
BCR-associated protein BAP31 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S71116
R:Adachi, T.; Schamel, W.W.A.; Kim, K.M.; Watanabe, T.; Becker, B.; Nielsen, P.J.; Reth,
EMBO J. 15, 1534-1541, 1996
A:Title: The specificity of association of the IgD molecule with the accessory proteins
A:Reference number: S71116; MUID:96203070; PMID:8612576
A:Accession: S71116
A:Molecule type: mRNA
A:Residues: 1-245 <ADA>
A:Cross-references: UNIPROT:Q61335; EMBL:X81816; NID:g1487985; PIDN:CAA57414.1; PID:g148
C:Genetics:
A:Gene: BAP31
A:Map position: X

Query Match      81.8%; Score 36; DB 2; Length 245;
Best Local Similarity 66.7%; Pred. No. 6.2;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 167 KLDIGNTGM 175

RESULT 3
T16972
Probable high affinity nitrate transporter protein - curled-leaved tobacco
C:Species: Nicotiana plumbaginifolia (curled-leaved tobacco)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16972
R:Quesada, A.; Krapp, A.; Trueman, L.; Daniel-Vedele, F.; Fernandez, E.; Forde, B.G.; Ca
Plant Mol. Biol. 34, 265-274, 1997
A:Title: PCR-identification of a Nicotiana plumbaginifolia cDNA homologous to the high a
A:Reference number: Z18625; MUID:97351560; PMID:9207842
A:Accession: T16972
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-530 <QUE>
A:Cross-references: UNIPROT:Q04431; EMBL:Y08210; NID:g2208959; PIDN:CAA69387.1; PID:g220
C:Genetics:
A:Gene: Nrt2; lnp
A:Superfamily: nitrate transporter component
C:Keywords: transmembrane protein

Query Match      79.5%; Score 35; DB 2; Length 530;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 97 KMDVGNAGV 105

RESULT 4
E96549
hypothetical protein F11M15.5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E96549
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

```

```

ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, R.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E96549
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1295 <STO>
A:Cross-references: UNIPROT:Q9SYC1; GB:A8005173; NID:g4836930; PIDN:AAD30632.1; GSPDB:GN
C:Genetics:
A:Gene: F11M15.5
A:Map position: 1

Query Match      79.5%; Score 35; DB 2; Length 1295;
Best Local Similarity 87.5%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LDVGNAEV 9
Db 470 LDVGNCEV 477

RESULT 5
A86192
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86192
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, R.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86192
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-174 <STO>
A:Cross-references: UNIPROT:Q9SE37; GB:A8005172; NID:g6850305; PIDN:AAF29382.1; GSPDB:GN
C:Genetics:
A:Map position: 1

Query Match      77.3%; Score 34; DB 2; Length 174;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAE 8
Db 158 KIDTGNAE 165

RESULT 6
D83626
transcription regulator PcaQ PA0152 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83626
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043

```

A:Accession: D83626  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-275 <STO>  
A:Cross-references: UNIPROT:Q916Y0; GB:AE004453; GB:AE004091; NID:g9945978; PIDN:AAG0354  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: pcaO; PA0152

Query Match 77.3%; Score 34; DB 2; Length 275;  
Best Local Similarity 66.7%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 224 LDVGNL 232

RESULT 7  
E71681  
isocitrate dehydrogenase (icd) RP265 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 09-Jul-2004  
C:Accession: E71681  
R:Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark, U.  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A:Reference number: A71630; MUID:99039499; PMID:9823893  
A:Accession: E71681  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-483 <AND>  
A:Cross-references: UNIPROT:Q92DR0; GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA1472  
A:Experimental source: strain Madrid E  
C:Genetics:  
A:Gene: icd; RP265  
C:Superfamily: isocitrate dehydrogenase (NADP)

Query Match 77.3%; Score 34; DB 2; Length 483;  
Best Local Similarity 66.7%; Pred. No. 35;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 399 KLDIGNFEL 407

RESULT 8  
A56374  
Glyceraldehyde-3-phosphate ferredoxin oxidoreductase (EC 1.-.-.-) - Pyrococcus furiosus  
N:Alternate names: GAPOR  
C:Species: Pyrococcus furiosus  
C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004  
C:Accession: A56374  
R:Mukund, S.; Adams, M.W.W.  
J. Biol. Chem. 270, 8389-8392, 1995  
A:Title: Glyceraldehyde-3-phosphate ferredoxin oxidoreductase, a novel tungsten-containing  
A:Reference number: A56374; MUID:95238315; PMID:7721730  
A:Accession: A56374  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-38 <MUR>  
A:Cross-references: UNIPROT:Q9UW12  
C:Comment: This 63k monomeric enzyme is sensitive to molecular oxygen. It contains pterin  
C:Keywords: iron; monomer; oxidoreductase

Query Match 75.0%; Score 33; DB 2; Length 38;  
Best Local Similarity 77.8%; Pred. No. 3.5;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 7 KLDVGKREV 15

RESULT 9  
H84414  
hypothetical protein Vng2652h [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: H84414  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
; Leithauser, B.; Keiler, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: H84414  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-443 <STO>  
A:Cross-references: UNIPROT:Q9HM87; GB:AE004437; NID:g10582033; PIDN:AAG20684.1; GSPDB:3  
C:Genetics:  
A:Gene: VNG2652H

Query Match 75.0%; Score 33; DB 2; Length 443;  
Best Local Similarity 66.7%; Pred. No. 52;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 171 KLDVQAEL 179

RESULT 10  
T03478  
probable DNA-directed DNA polymerase (EC 2.7.7.7) III alpha chain - Rhodobacter capsulatus  
N:Alternate names: DNA polymerase III alpha chain  
C:Species: Rhodobacter capsulatus  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C:Accession: T03478  
R:Vicek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.  
Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997  
A:Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1003  
A:Reference number: Z14955; MUID:97404404; PMID:9256491  
A:Accession: T03478  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-704 <VLC>  
A:Cross-references: UNIPROT:O68045; EMBL:AF010496; NID:g3128256; PIDN:AAC16131.1; PID:g3  
C:Genetics:  
A:Map position: 1  
C:Keywords: DNA replication; nucleotidyltransferase

Query Match 75.0%; Score 33; DB 2; Length 704;  
Best Local Similarity 87.5%; Pred. No. 86;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LDVGNAEV 9  
Db 428 LDVQAEL 435

RESULT 11  
D84145  
hypothetical protein BH3964 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C:Accession: D84145  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: D84145  
A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-65 <STO>  
 C:Cross-references: UNIPROT:Q9K5X3; GB:AF001520; GB:BA000004; NID:g10176401; PIDN:BA076  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: BH3964

Query Match 72.7%; Score 32; DB 2; Length 65;  
 Best Local Similarity 75.0%; Pred. No. 10;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAE 8  
 |||:||||  
 Db 26 KLDVGNAE 33

RESULT 12  
 AC0061  
 Pyridoxal phosphate biosynthetic protein [imported] - *Yersinia pestis* (strain CO92)  
 C:Species: *Yersinia pestis*  
 C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AC0061  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Hoiden, M.T.G.; Prentice, M.B.;  
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AC0061  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-331 <KUR>  
 A:Cross-references: UNIPROT:P58719; GB:AL590842; PIDN:CAC89350.1; PID:g15978587; GSPDB:G  
 C:Genetics:  
 A:Gene: pdxA  
 C:Superfamily: pdxA protein

Query Match 72.7%; Score 32; DB 2; Length 331;  
 Best Local Similarity 85.7%; Pred. No. 61;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNA 7  
 |||:||||  
 Db 89 KLDVGNS 95

RESULT 13  
 T45116  
 Phosphopyruvate hydratase (EC 4.2.1.11) [imported] - fission yeast (*Schizosaccharomyces*  
 C:Species: *Schizosaccharomyces pombe*  
 C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000  
 C:Accession: T45116  
 R:Park, S.K.  
 unpublished results 1994, cited by EMBL  
 A:Reference number: 222916  
 A:Accession: T45116  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: mRNA  
 A:Residues: 1-446 <PAR>  
 A:Cross-references: EMBL:L37084; PIDN:AAA51399.1  
 C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 72.7%; Score 32; DB 2; Length 446;  
 Best Local Similarity 62.5%; Pred. No. 85;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDVGNAEV 9  
 |||:||||  
 Db 118 LDVGNsqI 125

RESULT 14  
 A97744

isocitrate dehydrogenase (NADP) (EC 1.1.1.42) - *Rickettsia conorii* (strain Malish 7)  
 C:Species: *Rickettsia conorii*  
 C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
 C:Accession: A97744  
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro  
 Science 293, 2093-2098, 2001  
 A:Title: Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia prowazekii*.  
 A:Reference number: A97700; MUID:21442074; PMID:11557893  
 A:Accession: A97744  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-483 <KUR>  
 A:Cross-references: UNIPROT:Q92IR7; GB:AE006914; PIDN:AAL02891.1; PID:g15619416; GSPDB:G  
 C:Genetics:  
 A:Gene: icd  
 C:Superfamily: isocitrate dehydrogenase (NADP)  
 C:Keywords: oxidoreductase

Query Match 72.7%; Score 32; DB 2; Length 483;  
 Best Local Similarity 66.7%; Pred. No. 92;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
 |||:||||  
 Db 399 KLDLGNFEL 407

RESULT 15  
 T31802  
 hypothetical protein F29G9.2 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T31802  
 R:Langston, Y.

submitted to the EMBL Data Library, July 1997  
 A:Description: The sequence of *C. elegans* cosmid F29G9.  
 A:Reference number: Z21088

A:Accession: T31802  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-525 <LAN>  
 A:Cross-references: UNIPROT:Q8IA85; EMBL:AF016440; PIDN:AAB65904.1; GSPDB:GN00023; CESP:  
 A:Experimental source: strain Bristol N2; clone F29G9

C:Genetics:  
 A:Gene: CESP:F29G9.2  
 A:Map position: 5  
 A:Introns: 52/1; 91/3; 135/3; 333/3; 429/1

Query Match 72.7%; Score 32; DB 2; Length 525;  
 Best Local Similarity 55.6%; Pred. No. 1e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
 :||:||||  
 Db 427 EIDIGNLEV 435

Search completed: November 18, 2005, 01:09:29  
 Job time : 12.0032 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:04:26 ; Search time 39.9677 Seconds  
(without alignments)  
115.311 Million cell updates/sec

Title: US-10-006-177-3

Perfect score: 44

Sequence: 1 KLDVGNREV 9

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	44	100.0	245	1 BA31_HUMAN	P51572 homo sapien
2	38	86.4	866	2 Q8U202	Q8U202 pyrococcus
3	36	81.8	244	1 BA31_MOUSE	Q61335 mus musculus
4	36	81.8	379	1 HISZ_THETN	Q8T880 thermococcus
5	36	81.8	445	2 Q9GL39	Q9GL39 bos taurus
6	36	81.8	498	1 NWT2_BOVIN	Q9N181 bos taurus
7	35	79.5	206	2 Q6Q4H2	Q6Q4H2 oreochromis
8	35	79.5	439	2 Q6D532	Q6D532 erwinia car
9	35	79.5	530	2 Q04431	Q04431 nicotiana p
10	35	79.5	530	2 Q84M28	Q84M28 nicotiana t
11	35	79.5	530	2 Q84M29	Q84M29 nicotiana t
12	35	79.5	530	2 Q8SQ66	Q8SQ66 lycopersico
13	35	79.5	531	2 Q9SQG7	Q9SQG7 lycopersico
14	35	79.5	531	2 Q93X79	Q93X79 lycopersico
15	35	79.5	535	2 Q8RVG5	Q8RVG5 daucus caro
16	35	79.5	1295	2 Q9SYC1	Q9SYC1 arabidopsis
17	34	77.3	174	2 Q94K20	Q94K20 arabidopsis
18	34	77.3	174	2 Q9SE37	Q9SE37 arabidopsis
19	34	77.3	275	2 Q916X0	Q916X0 pseudomonas
20	34	77.3	277	2 Q21961	Q21961 bacterioph
21	34	77.3	386	2 Q6K6R8	Q6K6R8 oryza sativ
22	34	77.3	483	1 IDH_RICPR	Q9ZDR0 rickettsia
23	34	77.3	483	2 Q68XA5	Q68XA5 rickettsia
24	34	77.3	491	2 Q8EVG0	Q8EVG0 aeromonas s
25	34	77.3	503	2 Q6EVG7	Q6EVG7 aeromonas s
26	34	77.3	507	2 Q6EVF9	Q6EVF9 aeromonas s
27	34	77.3	675	2 Q7Y5E0	Q7Y5E0 bacterioph
28	34	77.3	688	2 Q6QGF0	Q6QGF0 bacterioph
29	33	75.0	38	2 Q9UW12	Q9UW12 pyrococcus
30	33	75.0	148	2 Q9WX84	Q9WX84 erysipeloth
31	33	75.0	188	2 Q8XJW8	Q8XJW8 clostridium

32	33	75.0	281	2	Q94E44	Q94E44 oryza sativ
33	33	75.0	283	2	Q7P2X3	Q7P2X3 fusobacteri
34	33	75.0	388	2	Q8SU07	Q8SU07 encephalito
35	33	75.0	411	2	Q68A45	Q68A45 raietonia s
36	33	75.0	416	2	Q8XR46	Q8XR46 raietonia s
37	33	75.0	443	2	Q9HM87	Q9HM87 halobacteri
38	33	75.0	460	1	NFHT_DROME	Q76464 d nitrilase
39	33	75.0	547	2	Q8SD67	Q8SD67 pseudomonas
40	33	75.0	653	2	Q93720	Q93720 pyrococcus
41	33	75.0	653	2	Q8U3K2	Q8U3K2 pyrococcus
42	33	75.0	704	1	D3JE_RHOCA	Q68045 rhodobacter
43	33	75.0	846	2	Q9VWN4	Q9VWN4 drosophila
44	33	75.0	1024	1	POPC_RALSO	Q9RBS2 raietonia s
45	33	75.0	1024	2	Q84IE6	Q84IE6 raietonia s

#### ALIGNMENTS

##### RESULT 1

BA31\_HUMAN  
ID BA31\_HUMAN STANDARD; PRT; 245 AA.  
AC P51572: Q13836; Q96CF0;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE B-cell receptor-associated protein 31 (BCR-associated protein Bap31)  
DE (p28 Bap31) (CDM protein) (6C6-AG tumor-associated antigen)  
DE (DXS1357E).  
GN Name=BCAP31; Synonyms=BAP31;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95104864; PubMed=7806238;  
RA Mosser J., Sarde C.O., Vicaire S., Yates J.R., Mandel J.L.;  
RT "A new human gene (DXS1357E) with ubiquitous expression, located in Xq28 adjacent to the adrenoleukodystrophy gene.";  
RL Genomics 22:469-471(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96300225; PubMed=8706661;  
RA Li E., Bestagno M., Burrone O.;  
RT "Molecular cloning and characterization of a transmembrane surface antigen in human cells.";  
RL Eur. J. Biochem. 238:631-638(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX TISSUE=B-cell, and promyelocytic leukemia;  
RC MEDLINE=96203070; PubMed=8612576;  
RA Adachi T., Schanel W.W.A., Kim K.-M., Watanabe T., Becker B., Nielsen P.J., Rath M.;  
RT "The specificity of association of the IgD molecule with the accessory proteins BAP31/BAP29 lies in the IgD transmembrane sequence.";  
RL EMBO J. 15:1534-1541(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Platzner M., Bauer D., Brenner V., Drescher B., Nyakatura G., Reichwald K., Sandoval N., Coy J., Kioschis P., Korn B., Poustka A., Rosenthal A.;  
RT Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,



OS Pyrococcus furiosus.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 ON NCBI\_TaxID=2261;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;  
 RT "The complete sequence of the Pyrococcus furiosus genome.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE010216; AAL81175.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR Pfam; PF00270; DEAD; 1.  
 DR SMART; PF00271; Helicase\_C; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 KW ATP-binding; Complete proteome; Helicase; Hydrolase.  
 SQ SEQUENCE 866 AA; 100236 MW; 12F80E7182AB3AA3 CRC64;  
 Query Match 86.4%; Score 38; DB 2; Length 866;  
 Best Local Similarity 66.7%; Pred No. 47;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KLDVGNABV 9  
 Db 796 KMDIGNAEL 804  
 ID BA31 MOUSE STANDARD; PRT; 244 AA.  
 AC Q61335; Q9D059; Q9DRG7;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE B-cell receptor-associated protein 31 (BCR-associated protein Bap31)  
 DE (p28 Bap31).  
 GN Name=Bcap31; Synonyms=Bap31;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.; SEQUENCE OF 1-10, INTERACTION WITH BCAP29 AND IGD,  
 RP AND TISSUE SPECIFICITY.  
 RC TISSUE=Plasmacytoma;  
 RX MEDLINE=96203070; PubMed=8612576;  
 RA Adachi T., Schamel W.W.A., Kim K.-M., Watanabe T., Becker B.,  
 RA Nielsen P.J., Reth M.;  
 RT "The specificity of association of the Igd molecule with the accessory  
 RT proteins BAP31/BAP29 lies in the Igd transmembrane sequence.";  
 RL EMBO J. 15:1534-1541(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo, and Small intestine;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi T., Bono H., Kondo S.,  
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusil V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gibbs C., Godzik A., Gough J.,  
 RA Grimmerd S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zvolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yaeunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast tumor;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 149-156; 167-179; 204-212; 213-219; 220-230 AND 231-241,  
 RP AND INTERACTION WITH BCAP29 AND IGD.  
 RX MEDLINE=9434992; PubMed=8070407;  
 RA Kim K.-M., Adachi T., Nielsen P.J., Terashima M., Lamers M.C.,  
 RA Kohler G., Reth M.;  
 RT "Two new proteins preferentially associated with membrane  
 RT immunoglobulin D.";  
 RL EMBO J. 13:3793-3800(1994).  
 CC -1- FUNCTION: May play a role in anterograde transport of membrane  
 CC proteins from the endoplasmic reticulum to the Golgi. May be  
 CC involved in CASP8-mediated apoptosis (By similarity).  
 CC -1- SUBUNIT: Homodimer and heterodimer with BCAP29. Binds CASP8  
 CC (isoform 9) as a complex containing BCAP31, BCAP29, BCL2 and/or  
 CC BCL2L1. Interacts with VAMP3, VAMP1 and membrane IGD  
 CC immunoglobulins. May interact with ACTG1 and non-muscle myosin II  
 CC (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
 CC reticulum. May shuttle between the ER and the intermediate  
 CC compartment/cis-Golgi complex (By similarity).  
 CC -1- TISSUE SPECIFICITY: Ubiquitous.  
 CC -1- PTM: Cleaved by CASP8 and other caspases (By similarity).  
 CC -1- SIMILARITY: Belongs to the BCAP29/BCAP31 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X81816; CAA57414.1; -



```
AC Q9N181;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Glycylpeptide N-tetradecanoyltransferase 2 (EC 2.3.1.97) (Peptide N-
DE myristoyltransferase 2) (Myristoyl-CoA:protein N-myristoyltransferase
DE 2) (NMT 2) (Type II N-myristoyltransferase).
GN Name=NMT2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Rundle D.R., Alvarez R.A., Anderson R.E.;
RT "Bovine retina type II N-myristoyltransferase.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adds a myristoyl group to the N-terminal glycine residue
CC of certain cellular and viral proteins (By similarity).
CC -!- CATALYTIC ACTIVITY: Tetradecanoyl-CoA + glycyl-peptide = CoA + N-
CC tetradecanoylglycyl-peptide.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the NMT family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF222687; AAF31456.1; -
DR HSSP; P30418; INMT.
DR InterPro; IPR000903; Myristoyl_trans.
DR Pfam; PF01233; NMT_1.
DR Pfam; PF02799; NMT_C; 1.
DR PROSITE; PS00975; NMT_1; 1.
DR PROSITE; PS00976; NMT_2; 1.
DR ACylttransferase; Transferase.
FT DOMAIN 46 56
SQ SEQUENCE 498 AA; 56802 MW; D3ECB34B98CA4D67 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 498;
Best Local Similarity 87.5%; Pred. No. 71;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDVGNAEV 9
Db 163 LDLGNAEV 170
||:|||||
||:|||||

RESULT 7
Q6Q4H2 PRELIMINARY; PRT; 206 AA.
AC Q6Q4H2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-1-antitrypsin (Fragment).
OS Oreochromis mossambicus (Mozambique tilapia) (Tilapia mossambica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8127;
RN [1]
RP SEQUENCE FROM N.A.
RA Lo M.J., Weng C.F.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the serpin family.
```

```
DR EMBL; AY559446; AAS67293.1; -.
DR HSSP; P01009; IKCT.
DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.
DR InterPro; IPR000215; Prot_inh_serpin.
DR Pfam; PF00079; Serpin; 1.
DR SMART; SM00093; SERPIN; 1.
KW Protease inhibitor; Serine protease inhibitor; Serpin.
FT NON_TER 206
FT NON_TER 206
SQ SEQUENCE 206 AA; 23538 MW; C824B9B6B2BA6705 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGN 7
Db 65 KLDVGN 71
|||||

RESULT 8
Q6D532 PRELIMINARY; PRT; 439 AA.
ID Q6D532;
AC Q6D532;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Periplasmic binding protein.
GN OrderedLocNames=ECA2210;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed-15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahnia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
DR EMBL; BX950851; CAG75111.1; -.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006059; SBP_bac_1.
DR Pfam; PF01547; SBP_bac_1; 1.
KW Complete proteome.
SQ SEQUENCE 439 AA; 48427 MW; C3965B4C71CFE8ED CRC64;

Query Match 79.5%; Score 35; DB 2; Length 439;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 411 KLDYGNAGEV 419
|||||

RESULT 9
Q04431 PRELIMINARY; PRT; 530 AA.
ID Q04431;
AC Q04431;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrate transporter.
GN Name=Nrt2; lnp;
```

```

OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97351560; PubMed=9207842; DOI=10.1023/A:1005872816881;
RA Quesada A., Krapp A., Trueman L., Daniel-Vedele F., Fernandez E.,
RA Forde B.G., Caboche M.;
RT "PCR-identification of a Nicotiana plumbaginifolia cDNA homologous to
RL the high affinity nitrate transporters of the crnA family.";
RL Plant Mol. Biol. 34:265-274(1997).
DR EMBL; Y08210; CAA69387.1; -.
DR PIR; T16972; T16972.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS0850; MFS; 1.
SQ SEQUENCE 530 AA; 57468 MW; 290CD622E0430992 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 530;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 97 KMDVGNAGV 105

RESULT 10
Q84MZ8
ID Q84MZ8 PRELIMINARY; PRT; 530 AA.
AC Q84MZ8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE High affinity nitrate transporter protein.
GN Name=nrt2.2;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Root;
RA Simon M., Vidmar J.J., Siddiqi Y.M., Kaiser B.N., Okamoto M.,
RA Gilbert M., Ellis D., Glass A.D.M.;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ557584; CAD89799.1; -.
SQ SEQUENCE 530 AA; 57726 MW; FEA575D4FF12C086 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 530;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 97 KMDVGNAGV 105

RESULT 11
Q84MZ9
ID Q84MZ9 PRELIMINARY; PRT; 530 AA.
AC Q84MZ9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE High affinity nitrate transporter protein.
GN Name=nrt2.1;
OS Nicotiana tabacum (Common tobacco).

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Root;
RA Simon M., Vidmar J.J., Siddiqi Y.M., Kaiser B.N., Okamoto M.,
RA Gilbert M., Ellis D., Glass A.D.M.;
RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ557583; CAD89798.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS0850; MFS; 1.
SQ SEQUENCE 530 AA; 57597 MW; D3BF4AD897A5746 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 530;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 97 KMDVGNAGV 105

RESULT 12
Q9SQG6
ID Q9SQG6 PRELIMINARY; PRT; 530 AA.
AC Q9SQG6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative high-affinity nitrate transporter.
GN Name=NRT2.1;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Root;
RA Ono F., Frommer W.B., von Wiren N.;
RT "Coordinated diurnal regulation of low- and high-affinity nitrate
RL transporters in tomato.";
RL Plant Biol. 2:17-23(2000).
DR EMBL; AF092655; AAF00053.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS0850; MFS; 1.
SQ SEQUENCE 530 AA; 57410 MW; 24422D21A614C2CA CRC64;

Query Match 79.5%; Score 35; DB 2; Length 530;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 97 KMDVGNAGV 105

RESULT 13
Q9SQG7
ID Q9SQG7 PRELIMINARY; PRT; 530 AA.
AC Q9SQG7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative high-affinity nitrate transporter.

```

GN Name=NRT2;2;  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Solanum.  
 OC NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Root;  
 RA Ono F., Frommer W.B., von Wiren N.;  
 RL "Coordinated diurnal regulation of low- and high-affinity nitrate  
 RT transporters in tomato";  
 RL Plant Biol. 2:17-23(2000).  
 DR EMBL; AF092654; AAF00054.1; --  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005215; P:transporter activity; IEA.  
 DR GO; GO:0006810; P:transporter; IEA.  
 DR InterPro; IPR007114; MFS.  
 DR PROSITE; PS50850; MFS; 1.  
 SQ SEQUENCE 530 AA; 57452 MW; E78BDF08C3E4CE33 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 530;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KLDVGNAEV 9  
 |:|||||  
 Db 97 KMDVGNAGV 105

RESULT 14  
 Q93X79 PRELIMINARY; PRT; 531 AA.  
 AC Q93X79;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Putative high-affinity nitrate transporter.  
 GN Name=NRT2;3;  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Solanum.  
 OC NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Root;  
 RX PubMed=12010476;  
 RA Hildebrandt U., Schmelzer E., Bothe H.;  
 RT "Expression of nitrate transporter genes in tomato colonized by an  
 RL arbuscular mycorrhizal fungus."  
 RL Physiol. Plantarum 115:125-136(2002).  
 DR EMBL; AY038800; AAK72402.1; --  
 SQ SEQUENCE 531 AA; 57549 MW; B7B863C9B52391D2 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 531;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KLDVGNAEV 9  
 |:|||||  
 Db 97 KMDVGNAGV 105

RESULT 15  
 Q8RVG5 PRELIMINARY; PRT; 535 AA.  
 ID Q8RVG5;  
 AC Q8RVG5;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE High-affinity nitrate transporter (Fragment).  
 OS Daucus carota (Carrot).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC campanulids; Apiales; Apiaceae; Apioidae; Scandiceae; Daucinae;  
 OC Daucus.  
 OC NCBI\_TaxID=4039;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mycorrhizae;  
 RA Lerner J.M., Clufo S., Wick R.L.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AV081213; AAL99362.1; --  
 FT NON\_TER 535 535  
 SQ SEQUENCE 535 AA; 58455 MW; 360C480C95B02998 CRC64;  
 Query Match 79.5%; Score 35; DB 2; Length 535;  
 Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 KLDVGNAEV 9  
 |:|||||  
 Db 94 KMDVGNAGV 102  
 Search completed: November 18, 2005, 01:07:59  
 Job time : 45.9677 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: November 18, 2005, 00:48:52 ; Search time 13.6452 Seconds  
(without alignments)  
49.237 Million cell updates/sec

Title: US-10-006-177-3

Perfect score: 44

Sequence: 1 KLDVGNAEV 9

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

- Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/pCTUS\_COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	246	4	US-09-180-167A-1
2	44	100.0	246	4	US-09-033-524B-1
3	36	81.8	245	4	US-09-180-167A-33
4	36	81.8	245	4	US-09-033-524B-33
5	35	79.5	63	4	US-09-180-167A-6
6	35	79.5	63	4	US-09-033-524B-6
7	35	79.5	656	4	US-09-903-540-9810
8	34	77.3	352	4	US-09-252-991A-19989
9	34	77.3	435	4	US-09-252-991A-28507
10	33	75.0	783	4	US-09-252-991A-32103
11	32	72.7	168	3	US-08-483-534A-2
12	32	72.7	215	4	US-09-949-016-7889
13	32	72.7	275	4	US-09-134-000C-4313
14	32	72.7	301	2	US-08-705-868-1
15	32	72.7	301	3	US-09-123-615-1
16	32	72.7	301	4	US-09-919-039A-132
17	32	72.7	393	4	US-09-248-796A-19608
18	32	72.7	456	4	US-09-634-238-276
19	32	72.7	492	4	US-09-949-016-10567
20	32	72.7	564	4	US-09-252-991A-23143
21	32	72.7	664	4	US-09-902-540-16365
22	32	72.7	973	4	US-09-107-532A-4810
23	32	72.7	1163	4	US-09-134-000C-5707
24	31	70.5	154	4	US-09-621-976-4212
25	31	70.5	314	4	US-09-489-039A-10377
26	31	70.5	327	4	US-09-949-016-7336
27	31	70.5	393	4	US-09-134-000C-5914

28	31	70.5	427	4	US-09-591-279A-25	Sequence 25, Appl
29	31	70.5	431	4	US-09-540-236-3536	Sequence 3536, Ap
30	31	70.5	461	3	US-09-355-115-7	Sequence 7, Appli
31	31	70.5	537	4	US-09-252-991A-20929	Sequence 20929, A
32	31	70.5	614	4	US-09-489-039A-12605	Sequence 12605, A
33	31	70.5	733	4	US-09-949-016-7651	Sequence 7651, Ap
34	31	70.5	925	4	US-09-252-991A-27057	Sequence 27057, A
35	31	70.5	1151	4	US-09-252-991A-21328	Sequence 21328, A
36	31	70.5	1270	4	US-10-101-464A-979	Sequence 979, App
37	31	70.5	1400	4	US-09-764-176-7	Sequence 7, Appli
38	30	68.2	62	4	US-09-107-433-3465	Sequence 3465, Ap
39	30	68.2	162	4	US-10-101-464A-611	Sequence 611, App
40	30	68.2	189	4	US-09-328-352-6722	Sequence 6722, Ap
41	30	68.2	197	4	US-09-270-767-42862	Sequence 42862, A
42	30	68.2	210	4	US-09-344-624-8	Sequence 8, Appli
43	30	68.2	300	3	US-09-134-001C-3422	Sequence 3422, Ap
44	30	68.2	315	4	US-09-462-846-7	Sequence 7, Appli
45	30	68.2	325	4	US-09-710-279-698	Sequence 698, App

## ALIGNMENTS

### RESULT 1

US-09-180-167A-1  
; Sequence 1, Application US/09180167A  
; Patent No. 6558950  
; GENERAL INFORMATION:  
; APPLICANT: Gordon C. Shore et al.  
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING  
; TITLE OF INVENTION: APOPTOSIS  
; FILE REFERENCE: 50013/004003  
; CURRENT APPLICATION NUMBER: US/09/180.167A  
; CURRENT FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: CA 2,198,988  
; PRIOR FILING DATE: 1997-03-03  
; PRIOR APPLICATION NUMBER: PCT/IB98/00706  
; PRIOR FILING DATE: 1998-03-02  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 246  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-180-167A-1

Query Match 100.0%; Score 44; DB 4; Length 246;  
Best Local Similarity 100.0%; Pred. No. 0.082;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 167 KLDVGNAEV 175

### RESULT 2

US-09-033-524B-1  
; Sequence 1, Application US/09033524B  
; Patent No. 6607880  
; GENERAL INFORMATION:  
; APPLICANT: Gordon C. Shore et al.  
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING  
; TITLE OF INVENTION: APOPTOSIS  
; FILE REFERENCE: 50013/004002  
; CURRENT APPLICATION NUMBER: US/09/033.524B  
; CURRENT FILING DATE: 1998-03-02  
; PRIOR APPLICATION NUMBER: CA 2,198,988  
; PRIOR FILING DATE: 1997-03-03  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 246  
; TYPE: PRT

```
; ORGANISM: Homo sapiens
US-09-033-524B-1

Query Match      100.0%; Score 44; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 167 KLDVGNAEV 175

RESULT 3
US-09-180-167A-33
; Sequence 33, Application US/09180167A
; Patent No. 6558950
; GENERAL INFORMATION:
; APPLICANT: Gordon C. Shore et al.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; FILE REFERENCE: 50013/004003
; CURRENT APPLICATION NUMBER: US/09/180,167A
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: CA 2,198,988
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: PCT/IB98/00706
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-180-167A-33

Query Match      81.8%; Score 36; DB 4; Length 245;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 167 KLDIGNTM 175

RESULT 4
US-09-033-524B-33
; Sequence 33, Application US/09033524B
; Patent No. 6607880
; GENERAL INFORMATION:
; APPLICANT: Gordon C. Shore et al.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; FILE REFERENCE: 50013/004002
; CURRENT APPLICATION NUMBER: US/09/033,524B
; CURRENT FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: CA 2,198,988
; PRIOR FILING DATE: 1997-03-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-033-524B-33

Query Match      81.8%; Score 36; DB 4; Length 245;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 167 KLDIGNTM 175

; ORGANISM: Homo sapiens
US-09-033-524B-1

Query Match      100.0%; Score 44; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 167 KLDVGNAEV 175

RESULT 5
US-09-180-167A-6
; Sequence 6, Application US/09180167A
; Patent No. 6558950
; GENERAL INFORMATION:
; APPLICANT: Gordon C. Shore et al.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; FILE REFERENCE: 50013/004003
; CURRENT APPLICATION NUMBER: US/09/180,167A
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: CA 2,198,988
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: PCT/IB98/00706
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-180-167A-6

Query Match      79.5%; Score 35; DB 4; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DVGNAEV 9
Db 1 DVGNAEV 7

RESULT 6
US-09-033-524B-6
; Sequence 6, Application US/09033524B
; Patent No. 6607880
; GENERAL INFORMATION:
; APPLICANT: Gordon C. Shore et al.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
; FILE REFERENCE: 50013/004002
; CURRENT APPLICATION NUMBER: US/09/033,524B
; CURRENT FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: CA 2,198,988
; PRIOR FILING DATE: 1997-03-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-033-524B-6

Query Match      79.5%; Score 35; DB 4; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DVGNAEV 9
Db 1 DVGNAEV 7

RESULT 7
US-09-902-540-9810
; Sequence 9810, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
```

FILE REFERENCE: 38-10(115849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883  
PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 9810  
LENGTH: 656  
TYPE: PRT  
ORGANISM: Myxococcus xanthus  
US-09-902-540-9810

Query Match 79.5%; Score 35; DB 4; Length 656;  
Best Local Similarity 77.8%; Pred. No. 26;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 479 KVDVGNAPV 487

## RESULT 8

US-09-252-991A-19989  
Sequence 19989, Application US/09252991A  
Patent No. 6551795

## GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 19989

LENGTH: 352

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-19989

Query Match 77.3%; Score 34; DB 4; Length 352;  
Best Local Similarity 87.5%; Pred. No. 21;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDVGNAEV 9  
Db 257 LDVGNAEV 264

## RESULT 9

US-09-252-991A-28507  
Sequence 28507, Application US/09252991A  
Patent No. 6551795

## GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 28507

LENGTH: 435

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28507

Query Match 77.3%; Score 34; DB 4; Length 435;  
Best Local Similarity 66.7%; Pred. No. 27;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 384 RLDVGNAGEL 392

## RESULT 10

US-09-252-991A-32103  
Sequence 32103, Application US/09252991A  
Patent No. 6551795

## GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 32103

LENGTH: 783

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-32103

Query Match 75.0%; Score 33; DB 4; Length 783;  
Best Local Similarity 55.6%; Pred. No. 91;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 383 RLDIGTAET 391

## RESULT 11

US-08-483-534A-2  
Sequence 2, Application US/08483534A  
Patent No. 6013483

## GENERAL INFORMATION:

APPLICANT: Coleman, Timothy A  
APPLICANT: Rosen, Craig  
TITLE OF INVENTION: Endothelial-Monocyte Activating  
TITLE OF INVENTION: Polypeptide III

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/483,534A

FILING DATE: 07 JUN 95

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-464 (PF206)

TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-483-534A-2

Query Match          72.7%; Score 32; DB 3; Length 168;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAE 8
Db 31 KIDVGEAE 38

RESULT 12
US-09-949-016-7889
; Sequence 7889, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7889
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7889

Query Match          72.7%; Score 32; DB 4; Length 215;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 18 KMDVGSKEV 26

RESULT 13
US-09-134-000C-4313
; Sequence 4313, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032798-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4313
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
```

```
US-09-134-000C-4313

Query Match          72.7%; Score 32; DB 4; Length 275;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 74 KLDVNNARI 82

RESULT 14
US-08-705-868-1
; Sequence 1, Application US/08705868
; Patent No. 5885798
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; APPLICANT: Au-Young, Janice
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,868
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0117 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 301 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-08-705-868-1

Query Match          72.7%; Score 32; DB 2; Length 301;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAB 8
Db 164 KIDVGEAE 171

RESULT 15
US-09-123-615-1
; Sequence 1, Application US/09123615
; Patent No. 6090377
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; APPLICANT: Au-Young, Janice
```

```

; APPLICANT: MURRY, LYNN E.
; TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/123,615
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,868
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0117 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 301 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
;
US-09-123-615-1

```

```

Query Match          72.7%; Score 32; DB 3; Length 301;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 KLDVGNAE 8
       |:|||||
Db      164 KIDVGEAE 171

```

Search completed: November 18, 2005, 09:35:55  
Job time : 15.6452 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 01:08:37 ; Search time 44.3226 Seconds  
(without alignments)  
84.961 Million cell updates/sec

Title: US-10-006-177-3

Perfect score: 44

Sequence: 1 KLDVGNAEV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	9	14 US-10-006-177-3	Sequence 3, Appli
2	44	100.0	87	16 US-10-408-765A-2451	Sequence 2451, Ap
3	44	100.0	123	18 US-10-450-763-58558	Sequence 58558, A
4	44	100.0	291	9 US-09-833-790-433	Sequence 433, App
5	44	100.0	313	15 US-10-094-749-3232	Sequence 3232, Ap
6	36	81.8	245	14 US-10-205-194-99	Sequence 99, Appl
7	36	81.8	715	15 US-10-424-599-209081	Sequence 209081,
8	34	77.3	250	16 US-10-425-115-235635	Sequence 235635,
9	34	77.3	275	15 US-10-282-122A-66155	Sequence 66155, A
10	33	75.0	95	16 US-10-767-701-39751	Sequence 39751, A
11	33	75.0	460	17 US-10-923-960-23	Sequence 23, Appl

12	33	75.0	460	20	US-11-097-143-15579	Sequence 15579, A
13	33	75.0	839	20	US-11-097-143-16410	Sequence 16410, A
14	33	75.0	891	15	US-10-307-817-590	Sequence 590, App
15	33	75.0	1108	16	US-10-437-963-149563	Sequence 149563,
16	32	72.7	73	16	US-10-425-115-227020	Sequence 227020, A
17	32	72.7	124	15	US-10-425-114-45099	Sequence 45099, A
18	32	72.7	162	15	US-10-425-114-69358	Sequence 69358, A
19	32	72.7	168	17	US-10-910-403-2	Sequence 2, Appli
20	32	72.7	171	15	US-10-282-122A-71353	Sequence 71353, A
21	32	72.7	178	9	US-09-813-718-6	Sequence 6, Appli
22	32	72.7	178	15	US-10-240-532-6	Sequence 6, Appli
23	32	72.7	178	16	US-10-240-527A-6	Sequence 6, Appli
24	32	72.7	216	9	US-09-925-302-482	Sequence 482, App
25	32	72.7	216	10	US-09-925-302-482	Sequence 482, App
26	32	72.7	259	16	US-10-437-963-159595	Sequence 159595,
27	32	72.7	275	15	US-10-424-599-167960	Sequence 167960,
28	32	72.7	277	16	US-10-425-115-296154	Sequence 296154,
29	32	72.7	301	10	US-09-919-039-132	Sequence 132, App
30	32	72.7	301	18	US-10-858-412-235	Sequence 235, App
31	32	72.7	326	16	US-10-425-115-202868	Sequence 202868,
32	32	72.7	341	16	US-10-437-963-131112	Sequence 131112,
33	32	72.7	351	18	US-10-858-412-234	Sequence 234, App
34	32	72.7	388	15	US-10-282-122A-67408	Sequence 67408, A
35	32	72.7	441	16	US-10-739-930-11049	Sequence 11049, A
36	32	72.7	456	15	US-10-264-213-179	Sequence 179, App
37	32	72.7	466	16	US-10-437-963-119990	Sequence 119990,
38	32	72.7	528	16	US-10-370-715B-372	Sequence 372, App
39	32	72.7	534	15	US-10-369-493-22005	Sequence 22005, A
40	32	72.7	536	9	US-09-813-718-2	Sequence 2, Appli
41	32	72.7	536	15	US-10-240-532-2	Sequence 2, Appli
42	32	72.7	536	16	US-10-240-527A-2	Sequence 2, Appli
43	32	72.7	556	15	US-10-389-647-625	Sequence 625, App
44	32	72.7	622	16	US-10-425-115-344123	Sequence 344123,
45	32	72.7	632	16	US-10-437-963-146725	Sequence 146725,

#### ALIGNMENTS

#### RESULT 1

US-10-006-177-3  
; Sequence 3, Application US/10006177  
; Publication No. US20030165513A1  
; GENERAL INFORMATION:  
; APPLICANT: Ramakrishna, Venky  
; APPLICANT: Philip, Ramila  
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treatm  
; FILE REFERENCE: 26747-35  
; CURRENT APPLICATION NUMBER: US/10/006.177  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Epitopic Peptide  
US-10-006-177-3

Query Match 100.0%; Score 44; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9

Db 1 KLDVGNAEV 9

```
RESULT 2
US-10-408-765A-2451
; Sequence 2451, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2451
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2451

Query Match      100.0%; Score 44; DB 16; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 27 KLDVGNAEV 35

RESULT 3
US-10-450-763-58558
; Sequence 58558, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 58558
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: (39)-(93)
; OTHER INFORMATION: Kinesin light chain repeat proteins domain identified by
; OTHER INFORMATION: eMATRIX, accession number BL01160B, p-value=5.653e-09, raw score
; OTHER INFORMATION: 19.54
US-10-450-763-58558

Query Match      100.0%; Score 44; DB 18; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 43 KLDVGNAEV 51

RESULT 4
US-09-833-790-433
; Sequence 433, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indirias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 433
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-433

Query Match      100.0%; Score 44; DB 9; Length 291;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9
Db 212 KLDVGNAEV 220

RESULT 5
US-10-094-749-3232
; Sequence 3232, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3232
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-3232
```



Query Match 100.0%; Score 44; DB 15; Length 313;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 234 KLDVGNAEV 242

## RESULT 6

US-10-205-194-99  
; Sequence 99, Application US/10205194  
; Publication No. US20030134301A1  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert Company  
; APPLICANT: Lee, Kevin  
; APPLICANT: Dixon, Alistair  
; APPLICANT: Brooksbank, Robert  
; APPLICANT: Pinnock, Robert  
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
; FILE REFERENCE: WL-A-018201  
; CURRENT APPLICATION NUMBER: US/10/205,194  
; CURRENT FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: GB 0118354.0  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 177  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 99  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: BAP31  
US-10-205-194-99

Query Match 81.8%; Score 36; DB 14; Length 245;  
Best Local Similarity 66.7%; Pred. No. 37;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 167 KLDIGNTEM 175

## RESULT 7

US-10-424-599-209081  
; Sequence 209081, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 209081  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1) .. (715)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_30828C.1.pep  
US-10-424-599-209081

Query Match 81.8%; Score 36; DB 15; Length 715;

Best Local Similarity 87.5%; Pred. No. 1.2e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDVGNAEV 9  
Db 127 VDVGNAEV 134

## RESULT 8

US-10-425-115-235635  
; Sequence 235635, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 235635  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_146484C.1.pep  
US-10-425-115-235635

Query Match 77.3%; Score 34; DB 16; Length 250;  
Best Local Similarity 66.7%; Pred. No. 1e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLDVGNAEV 9  
Db 86 ELVDGNSEL 94

## RESULT 9

US-10-282-122A-66155  
; Sequence 66155, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23

```
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66155
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-10-282-122A-66155

Query Match      77.3%; Score 34; DB 15; Length 275;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 KLDVGNAEV 9
Db      224 RLDVGNEL 232

RESULT 10
US-10-767-701-39751
; Sequence 39751, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 39751
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C29032_1.pep
; US-10-767-701-39751

Query Match      75.0%; Score 33; DB 16; Length 95;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 KLDVGNAEV 9
Db      67 KLEPGNAEI 75

RESULT 11
US-10-923-960-23
; Sequence 23, Application US/10923960
; Publication No. US20050019890A1
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; TITLE OF INVENTION: Nitrilase Homologs
; FILE REFERENCE: TJU-2510
; CURRENT APPLICATION NUMBER: US/10/923,960
; CURRENT FILING DATE: 2004-08-23
; PRIOR APPLICATION NUMBER: US/09/357,675
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/093,350
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 23
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-10-923-960-23

Query Match      75.0%; Score 33; DB 17; Length 460;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 KLDVGNAEV 9
Db      269 ELDIGTAEV 277

RESULT 12
US-11-097-143-15579
; Sequence 15579, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15579
; LENGTH: 460
; TYPE: PRT
; ORGANISM: DROSOPHILA
; US-11-097-143-15579

Query Match      75.0%; Score 33; DB 20; Length 460;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 KLDVGNAEV 9
Db      269 ELDIGTAEV 277

RESULT 13
US-11-097-143-16410
; Sequence 16410, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
```

```
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16410
; LENGTH: 839
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-16410
```

```
Query Match 75.0%; Score 33; DB 20; Length 839;
Best Local Similarity 55.6%; Pred. No. 6.2e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 KLDVGNAEV 9
Db 65 EVDIGNSEV 73
:::|||||
```

```
RESULT 14
US-10-307-817-590
; Sequence 590, Application US/10307817
; Publication No. US20040058338A1
; GENERAL INFORMATION:
; APPLICANT: Agee et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-502C
; CURRENT APPLICATION NUMBER: US/10/307,817
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 590
; LENGTH: 991
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-307-817-590
```

```
Query Match 75.0%; Score 33; DB 15; Length 991;
Best Local Similarity 66.7%; Pred. No. 7.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 KLDVGNAEV 9
Db 452 KQVGNTEV 460
|:|||||
```

```
RESULT 15
US-10-437-963-149563
; Sequence 149563, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
```

```
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149563
; LENGTH: 1108
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1108)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_49884C.1.pap
US-10-437-963-149563
```

```
Query Match 75.0%; Score 33; DB 16; Length 1108;
Best Local Similarity 66.7%; Pred. No. 8.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 KLDVGNAEV 9
Db 1082 KLDGNGVEI 1090
||| |||
```

```
Search completed: November 18, 2005, 10:44:32
Job time : 47.3226 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:03:06 ; Search time 48.7742 Seconds  
(without alignments)  
71.366 Million cell updates/sec

Title: US-10-006-177-5  
Perfect score: 47  
Sequence: 1 ALMEQQHV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:.\*  
1: geneseqp1980s:.\*  
2: geneseqp1990s:.\*  
3: geneseqp2000s:.\*  
4: geneseqp2001s:.\*  
5: geneseqp2002s:.\*  
6: geneseqp2003as:.\*  
7: geneseqp2003bs:.\*  
8: geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	9	ABG32305	Abg32305 HLA-A2 as
2	47	100.0	768	Aar27684	Aar27684 Human bet
3	47	100.0	769	Abg96351	Abg96351 Human ova
4	47	100.0	769	Abu56723	Abu56723 Lung canc
5	47	100.0	769	Abd80492	Abd80492 Ovarian c
6	47	100.0	769	Adn39527	Adn39527 Cancer/an
7	47	100.0	769	Adn39162	Adn39162 Cancer/an
8	47	100.0	769	Adn39507	Adn39507 Cancer/an
9	47	100.0	769	Adm86616	Adm86616 Human int
10	35	74.5	353	Abm68130	Abm68130 Photorhab
11	34	72.3	153	Aau15919	Aau15919 Human nov
12	34	72.3	153	Abu54988	Abu54988 Human nov
13	34	72.3	294	Adj68967	Adj68967 Human hea
14	34	72.3	302	Aab41960	Aab41960 Human ORF
15	34	72.3	319	Adq67763	Adq67763 Novel hum
16	33	70.2	229	Abb94775	Abb94775 Human pro
17	33	70.2	229	Abb81876	Abb81876 Double st
18	33	70.2	229	Aau81227	Aau81227 Human lun
19	33	70.2	268	Ado02317	Ado02317 Thalecres
20	33	70.2	302	Abb58029	Abb58029 Drosophil
21	33	70.2	311	Abro1655	Abro1655 Human G p
22	33	70.2	619	Adq67755	Adq67755 Novel hum
23	33	70.2	754	Adn27144	Adn27144 Bacterial
24	33	70.2	775	Aau93161	Aau93161 Arabidops
25	33	70.2	881	Aus43822	Aus43822 Bacterial

26	33	70.2	1318	8	ADQ66591	Novel hum
27	33	70.2	2359	5	ABB91251	Herbicida
28	32	68.1	152	7	ADF07220	Bacterial
29	32	68.1	191	7	ABO66089	Abob6089 Klebsiell
30	32	68.1	241	8	ADS42684	Bacterial
31	32	68.1	365	6	ABP78558	N. gonorr
32	32	68.1	455	6	ABM68247	Photorhab
33	32	68.1	591	2	AAZ22330	Linoleate
34	32	68.1	591	4	AAAB66306	L reuteri
35	32	68.1	591	8	ADJ34085	L. reuter
36	32	68.1	592	8	ADJ34079	L. reuter
37	32	68.1	634	4	ABB71624	Drosophil
38	32	68.1	1400	4	ABB59555	Drosophil
39	32	68.1	1400	7	ADK11344	Drosophil
40	32	68.1	2109	8	ADI26360	Aspergill
41	31	66.0	46	4	AAU14742	Novel bon
42	31	66.0	79	7	ADC88095	Ribosomal
43	31	66.0	164	7	ADF07913	Bacterial
44	31	66.0	171	6	ABU42129	Protein e
45	31	66.0	171	6	ABU42078	Protein e

## ALIGNMENTS

RESULT 1  
ABG32305  
ID ABG32305 standard; peptide; 9 AA.  
XX  
AC ABG32305;  
XX  
DT 05-NOV-2002 (first entry)  
XX  
DE HLA-A2 associated immunogenic peptide from human integrin beta-8.  
XX  
KW Human; immunogen; epitope; HLA-A1; human leukocyte antigen; CTL; cytotoxic Y lymphocyte; cytostatic; cancer; colorectal carcinoma; ovarian carcinoma; lung carcinoma; prostate carcinoma; vaccine; tumour;  
KW HLA-2; passive immunotherapy; integrin beta-8 subunit precursor.  
OS Homo sapiens.  
XX  
PN WO200246416-A2.  
XX  
PD 13-JUN-2002.  
XX  
PF 04-DEC-2001; 2001WO-US047290.  
XX  
PR 04-DEC-2000; 2000US-0251022P.  
XX 20-DEC-2000; 2000US-0256824P.  
(ARGO-) ARGONEX INC.  
Ramakrishna V, Ross M, Philip R;  
WPI; 2002-619021/66.  
New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte, and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.  
Claim 1; Page 50; 60pp; English.

The invention relates to an immunogen comprising an isolated polypeptide whose amino acid sequence comprises an epitopic peptide, does not include MAGE 4 or MFG-E8 proteins, or consists of MAGE D protein or its immunologically active fragment. Also included are a polynucleotide encoding the immunogen or its complement, a vector comprising the polynucleotide, a mammalian cell comprising the vector and expressing the polynucleotide, a vaccine composition comprising the immunogen and an antibody specific for the immunogen. The immunogen is useful for inducing a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is useful for inducing a CTL response when administered to a subject. A



CC characterising cancer, in detecting the presence of cancer as early as  
 CC possible, and the recurrence of ovarian cancer. The method may also be of  
 CC particular use with patients having an enhanced risk of developing  
 CC ovarian cancer (e.g. patients having a familial history of ovarian  
 CC cancer). The cancer markers may be used in the management and treatment  
 CC of e.g. brain and central nervous system disorders (e.g. bacterial and  
 CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain  
 CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),  
 CC inflammations (e.g. bacterial or viral meningitis or encephalitis),  
 CC testicular disorders (e.g. nontuberculous granulomatous orchitis),  
 CC connective tissue disorders, or heart disorders (e.g. ischaemic heart  
 CC disease or atherosclerosis). The compositions and methods may also be  
 CC used in assessing the histological type of neoplasm associated with  
 CC ovarian cancer, monitoring the progression of ovarian cancer, determining  
 CC whether ovarian cancer has metastasized or is likely to metastasize,  
 CC selecting a composition for inhibiting ovarian cancer, assessing the  
 CC ovarian carcinogenic potential of a compound, or inhibiting ovarian  
 CC cancer or at risk of developing ovarian cancer. The present amino acid  
 CC sequence represents one of the ovarian cancer markers described in the  
 CC invention  
 XX  
 SQ Sequence 769 AA;

Query Match 100.0%; Score 47; DB 5; Length 769;  
 Best Local Similarity 100.0%; Pred. No. 1.6; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQQHYV 9  
 |||||  
 Db 662 ALMEQQHYV 670

RESULT 4  
 ABUS6723  
 ID ABUS6723 standard; protein; 769 AA.

XX AC ABUS6723;  
 XX  
 DT 02-APR-2003 (first entry)  
 XX  
 DE Lung cancer-associated polypeptide #316.  
 XX  
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;  
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
 XX  
 OS Unidentified.

XX WO200286443-A2.  
 PN

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

PR 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-0334370P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

DR N-PSDB; ABX76452.

XX  
 PT Detecting a lung cancer-associated transcript in a cell from a patient  
 PT for treating lung cancer, by contacting a biological sample from the  
 PT patient with a polynucleotide that exhibits increased or decreased

PT expression in lung cancer.

XX Claim 27; Page 433; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated  
 CC transcript in a cell from a patient, comprising contacting a biological  
 CC sample from the patient with a polynucleotide that selectively hybridizes  
 CC to a sequence that is at least 80 % identical to a gene that exhibits  
 CC increased or decreased expression in lung cancer samples. Lung cancer-  
 CC associated polynucleotides and polypeptides are used for identifying a  
 CC compound that modulates a lung cancer-associated polypeptide, for  
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung  
 CC cancer in a patient and for treating a mammal having lung cancer by  
 CC administering a modulatory compound identified. The methods are useful  
 CC for treating lung cancer, such as small cell lung cancer, non-small cell  
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful  
 CC for diagnostic purposes and as targets for screening for therapeutic  
 CC compounds that modulate lung cancer, such as antibodies. Sequences  
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the  
 CC invention

XX Sequence 769 AA;

Query Match 100.0%; Score 47; DB 6; Length 769;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQQHYV 9  
 |||||  
 Db 662 ALMEQQHYV 670

RESULT 5

ADB80492

ID ADB80492 standard; protein; 769 AA.

XX AC ADB80492;

XX 04-DEC-2003 (first entry)

DE Ovarian cancer-associated protein #28.

XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;

XX post-operative chemotherapy; radiation therapy; tumour prognosis;

XX pre-cancerous lesion detection.

XX Homo sapiens.

XX WO2002102235-A2.

XX 27-DEC-2002.

XX 18-JUN-2002; 2002WO-US019297.

XX 18-JUN-2001; 2001US-0299234P.

PR 27-AUG-2001; 2001US-0315287P.

PR 05-SEP-2001; 2001US-0317544P.

PR 13-NOV-2001; 2001US-0350666P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC;

XX WPI; 2003-167431/16.

DR N-PSDB; ADB80491.

XX  
 PT Detecting an ovarian cancer-associated transcript in a cell from a  
 PT patient, comprises contacting a biological sample from the patient with a  
 PT polynucleotide that hybridizes to an ovarian cancer gene.

```

XX PS Claim 13; Page 293; 332pp; English.
XX CC The invention relates to a method of detecting an ovarian cancer-
XX CC associated transcript in a cell from a patient, by contacting a
XX CC biological sample from the patient with a polynucleotide that selectively
XX CC hybridizes to a sequence at least 80% identical to any of one of 80
XX CC nucleic acid sequences given in the specification. The method is useful
XX CC in diagnosing ovarian cancer and in identifying and using agents and/or
XX CC targets that inhibit ovarian cancer. The nucleic acid molecule,
XX CC polypeptide and the antibody may also be used in detecting ovarian
XX CC cancers, monitoring and early detection of relapse following treatment,
XX CC monitoring response to therapy, selecting patients for post-operative
XX CC chemotherapy or radiation therapy, in selecting mode of therapy,
XX CC determining tumour prognosis, early detection of pre-cancerous lesions,
XX CC and as vaccines. This sequence corresponds to one of the proteins used
XX CC for the detection method of the invention.
XX SQ Sequence 769 AA;

Query Match 100.0%; Score 47; DB 7; Length 769;
Best Local Similarity 100.0%; Pred. NO. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQHHYV 9
Db 662 ALMEQHHYV 670

RESULT 6
ADN39527
ID ADN39527 standard; protein; 769 AA.
XX AC ADN39527;
XX CC
XX DT 17-JUN-2004 (first entry)
XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A127.
XX KW Human; differential expression; cancer; angiogenic disorder;
XX KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
XX KW inflammatory disease; autoimmune disease;
XX KW retinal neovascularisation syndrome; scarring; uterine fibroid;
XX KW detection; diagnosis; prognosis; drug screening; drug targeting;
XX KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
XX KW vulnerary; gene therapy; vaccine.
XX OS Homo sapiens.
XX PN WO2003042661-A2.
XX PD 22-MAY-2003.
XX PF 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 21-NOV-2001; 2001US-0332464P.
XX PR 29-NOV-2001; 2001US-0334393P.
XX PR 03-DEC-2001; 2001US-0335394P.
XX PR 14-DEC-2001; 2001US-0340376P.
XX PR 08-JAN-2002; 2002US-0347211P.
XX PR 10-JAN-2002; 2002US-0347349P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PR 08-FEB-2002; 2002US-0355250P.
XX PR 13-FEB-2002; 2002US-0356714P.
XX PR 20-FEB-2002; 2002US-0359077P.
XX PR 29-MAR-2002; 2002US-0368809P.
XX PR 04-APR-2002; 2002US-0370110P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PR 05-JUN-2002; 2002US-0386614P.
XX PR 16-JUL-2002; 2002US-0396839P.
XX PR 22-JUL-2002; 2002US-0397775P.
XX PR 22-JUL-2002; 2002US-0397845P.
XX PR 09-SEP-2002; 2002US-0409450P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
XX PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WI: 2003-468649/44.
XX DR N-PSDB; ADN39526.
XX CC Determining the presence or absence of a pathological cell in a patient,
XX CC useful for diagnosing, prognosing or treating cancer, comprises detecting
XX CC a nucleic acid in a biological sample.
XX CC Claim 12; SEQ ID NO A127; 1385pp; English.
XX CC The invention relates to nucleic acids and proteins (ADN38583-ADN40064)
XX CC whose expression is upregulated or downregulated in specific cancers or
XX CC other diseases such as angiogenic or fibrotic disorders, and to methods
XX CC of determining the presence or absence of a pathological cell in a
XX CC patient by detecting a nucleic acid at least 80% identical to those of
XX CC the invention or by detecting a polypeptide of the invention. The
XX CC invention also relates to expression vectors and host cells comprising a
XX CC nucleic acid of the invention; antibodies which specifically bind a
XX CC polypeptide of the invention; use of such antibodies for drug targeting;
XX CC and methods of screening for modulators of activity or expression of the
XX CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX CC antibodies and methods are useful for diagnosing, prognosing and treating
XX CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX CC neovascularisation syndromes, scarring and uterine fibroids. They may
XX CC also be useful in wound healing and in contraception. The present
XX CC sequence represents a polypeptide of the invention.
XX SQ Sequence 769 AA;

Query Match 100.0%; Score 47; DB 7; Length 769;
Best Local Similarity 100.0%; Pred. NO. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQHHYV 9
Db 662 ALMEQHHYV 670

RESULT 7
ADN39162
ID ADN39162 standard; protein; 769 AA.
XX AC ADN39162;
XX CC
XX DT 17-JUN-2004 (first entry)
XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:480.
XX KW Human; differential expression; cancer; angiogenic disorder;
XX KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
XX KW inflammatory disease; autoimmune disease;
XX KW retinal neovascularisation syndrome; scarring; uterine fibroid;
XX KW detection; diagnosis; prognosis; drug screening; drug targeting;
XX KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
XX KW vulnerary; gene therapy; vaccine.
XX OS Homo sapiens.
XX PN WO2003042661-A2.
XX PD 22-MAY-2003.
XX PF 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 21-NOV-2001; 2001US-0332464P.
XX PR 29-NOV-2001; 2001US-0334393P.

```



PR 03-DEC-2001; 2001US-0335394P.  
 PR 14-DEC-2001; 2001US-0340376P.  
 PR 08-JAN-2002; 2002US-0347211P.  
 PR 10-JAN-2002; 2002US-0347349P.  
 PR 08-FEB-2002; 2002US-0355250P.  
 PR 13-FEB-2002; 2002US-0356714P.  
 PR 20-FEB-2002; 2002US-0359077P.  
 PR 29-MAR-2002; 2002US-036809P.  
 PR 04-APR-2002; 2002US-0370110P.  
 PR 12-APR-2002; 2002US-0372246P.  
 PR 05-JUN-2002; 2002US-0386614P.  
 PR 16-JUL-2002; 2002US-0396839P.  
 PR 22-JUL-2002; 2002US-039775P.  
 PR 09-SEP-2002; 2002US-0409450P.  
 XX PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;  
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;  
 XX DR WPI; 2003-468649/44.  
 DR N-PSDB; ADN39161.  
 XX PT Determining the presence or absence of a pathological cell in a patient,  
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting  
 PT a nucleic acid in a biological sample.  
 XX FS Claim 12; SEQ ID NO 480; 1385pp; English.  
 XX CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)  
 CC whose expression is upregulated or downregulated in specific cancers or  
 CC other diseases such as angiogenic or fibrotic disorders, and to methods  
 CC of determining the presence or absence of a pathological cell in a  
 CC patient by detecting a nucleic acid at least 80% identical to those of  
 CC the invention or by detecting a polypeptide of the invention. The  
 CC invention also relates to expression vectors and host cells comprising a  
 CC nucleic acid of the invention; antibodies which specifically bind a  
 CC polypeptide of the invention; use of such antibodies for drug targeting;  
 CC and methods of screening for modulators of activity or expression of the  
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,  
 CC antibodies and methods are useful for diagnosing, prognosing and treating  
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,  
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
 CC neovascularisation syndromes, scarring and uterine fibroids. They may  
 CC also be useful in wound healing and in contraception. The present  
 CC sequence represents a polypeptide of the invention.  
 XX SQ Sequence 769 AA;  
 Query Match 100.0%; Score 47; DB 7; Length 769;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALMEQQHYV 9  
 DB 662 ALMEQQHYV 670  
 RESULT 8  
 ADN39507  
 ID ADN39507 standard; protein; 769 AA.  
 XX AC ADN39507;  
 XX DT 17-JUN-2004 (first entry)  
 XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:ADN39507.  
 XX KW Human; differential expression; cancer; angiogenic disorder;  
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
 KW inflammatory disease; autoimmune disease;  
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;

KW detection; diagnosis; prognosis; drug screening; drug targeting;  
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;  
 XX vulnery; gene therapy; vaccine.  
 OS Homo sapiens.  
 XX WO2003042661-A2.  
 XX PD 22-MAY-2003.  
 XX PF 13-NOV-2002; 2002WO-US036810.  
 XX PR 13-NOV-2001; 2001US-0350666P.  
 PR 21-NOV-2001; 2001US-0332464P.  
 PR 29-NOV-2001; 2001US-0334393P.  
 PR 03-DEC-2001; 2001US-0335394P.  
 PR 14-DEC-2001; 2001US-0340376P.  
 PR 08-JAN-2002; 2002US-0347211P.  
 PR 10-JAN-2002; 2002US-0347349P.  
 PR 08-FEB-2002; 2002US-0355250P.  
 PR 13-FEB-2002; 2002US-0356714P.  
 PR 20-FEB-2002; 2002US-0359077P.  
 PR 29-MAR-2002; 2002US-036809P.  
 PR 04-APR-2002; 2002US-0370110P.  
 PR 12-APR-2002; 2002US-0372246P.  
 PR 05-JUN-2002; 2002US-0386614P.  
 PR 16-JUL-2002; 2002US-0396839P.  
 PR 22-JUL-2002; 2002US-039775P.  
 PR 09-SEP-2002; 2002US-0409450P.  
 XX (EOSB-) EOS BIOTECHNOLOGY INC.  
 PA Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;  
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;  
 XX WPI; 2003-468649/44.  
 DR N-PSDB; ADN39506.  
 XX PT Determining the presence or absence of a pathological cell in a patient,  
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting  
 PT a nucleic acid in a biological sample.  
 XX FS Claim 12; SEQ ID NO A107; 1385pp; English.  
 XX CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)  
 CC whose expression is upregulated or downregulated in specific cancers or  
 CC other diseases such as angiogenic or fibrotic disorders, and to methods  
 CC of determining the presence or absence of a pathological cell in a  
 CC patient by detecting a nucleic acid at least 80% identical to those of  
 CC the invention or by detecting a polypeptide of the invention. The  
 CC invention also relates to expression vectors and host cells comprising a  
 CC nucleic acid of the invention; antibodies which specifically bind a  
 CC polypeptide of the invention; use of such antibodies for drug targeting;  
 CC and methods of screening for modulators of activity or expression of the  
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,  
 CC antibodies and methods are useful for diagnosing, prognosing and treating  
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,  
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
 CC neovascularisation syndromes, scarring and uterine fibroids. They may  
 CC also be useful in wound healing and in contraception. The present  
 CC sequence represents a polypeptide of the invention.  
 XX SQ Sequence 769 AA;  
 Query Match 100.0%; Score 47; DB 7; Length 769;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALMEQQHYV 9  
 DB 662 ALMEQQHYV 670  
 RESULT 8  
 ADN39507  
 ID ADN39507 standard; protein; 769 AA.  
 XX AC ADN39507;  
 XX DT 17-JUN-2004 (first entry)  
 XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:ADN39507.  
 XX KW Human; differential expression; cancer; angiogenic disorder;  
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
 KW inflammatory disease; autoimmune disease;  
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;

RESULT 9  
ADE86616  
ID ADE86616 standard; protein; 769 AA.  
XX  
XX  
AC ADE86616;  
XX  
XX 29-JAN-2004 (first entry)  
XX  
XX Human integrin beta-8 subunit.  
XX  
XX human; secreted protein; cancer; liver disorder; hepatitis;  
XX KW neural disorder; Alzheimer's disease; integrin beta-8 subunit.  
XX  
XX Homo sapiens.  
XX  
XX US2003129685-A1.  
XX  
XX 10-JUL-2003.  
XX  
XX 18-APR-2001; 2001US-00836353.  
XX  
XX 28-OCT-1998; 98US-0105971P.  
XX PR 27-OCT-1999; 99WO-US025031.  
XX PR 19-APR-2000; 2000US-0198407P.  
XX  
XX (NIJJ//) NI J.  
XX PA (YOUN//) YOUNG P E.  
XX PA (KENN//) KENNY J J.  
XX PA (OLSE//) OLSEN H S.  
XX PA (MOOR//) MOORE P A.  
XX PA (WEIY//) WEI Y.  
XX PA (GREE//) GREENE J M.  
XX PA (RUBE//) RUBEN S M.  
XX  
XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
XX PI Ruben SM;  
XX  
XX WPI; 2004-020335/02.  
XX  
XX New nucleic acid molecule, useful for preparing a medicament for  
XX PT preventing, treating or ameliorating a medical condition e.g. cancer,  
XX PT liver disorders or neural disorders.  
XX  
XX Disclosure; SEQ ID NO 67; 380pp; English.  
XX  
XX The invention relates to an isolated nucleic acid sequence, or its  
XX CC allelic variant, a fragment of the cDNA sequence, or its fragment,  
XX CC domain, epitope or species homologue. The nucleic acid is useful for  
XX CC preparing a medicament for preventing, treating or ameliorating a medical  
XX CC condition e.g., cancer, liver disorders such as hepatitis or neural  
XX CC disorders such as Alzheimer's disease. The present sequence represents  
XX CC the amino acid sequence of an integrin beta-8 subunit.  
XX  
XX Sequence 769 AA;  
SQ  
Query Match 100.0%; Score 47; DB 8; Length 769;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ALMEQQHYV 9  
Db 662 ALMEQQHYV 670  
RESULT 10  
ABM68130  
ID ABM68130 standard; protein; 353 AA.  
XX  
XX ABM68130;  
XX  
XX 20-NOV-2003 (first entry)  
XX  
XX

DE Photorhabdus luminescens protein sequence #1227.  
XX  
KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
KW detection; food; gene expression; plant; animal; microorganism; toxin;  
KW antibiotic; biopesticide; virulence factor; disease model; plague;  
KW whooping cough.  
XX  
OS Photorhabdus luminescens.  
XX  
XX WO200294867-A2.  
XX  
XX 28-NOV-2002.  
XX  
XX 07-FEB-2002; 2002WO-IB003040.  
XX  
XX 07-FEB-2001; 2001FR-00001659.  
XX PR  
XX (INSP ) INST PASTEUR.  
XX PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
XX PI Buchrieser C;  
XX  
XX WPI; 2003-148459/14.  
XX  
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
XX PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
XX  
XX Claim 2; SEQ ID NO 1227; 1205pp; French.  
XX  
XX The invention relates to the isolation of genes and their encoded  
XX CC proteins from Photorhabdus luminescens. The isolated sequences are  
XX CC sources of probes and primers for detecting the genome of P. luminescens  
XX CC and related species; to study polymorphisms; for gene analysis and for  
XX CC detection/amplification of the genes. Antibodies (Ab) raised against the  
XX CC polypeptides encoded by the genes are used for detection/identification  
XX CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that  
XX CC carry a gene-containing vector are used to select compounds that  
XX CC modulate, regulate, induce or inhibit expression of the genes in plants,  
XX CC animals or microorganisms other than P. luminescens and are able to alter  
XX CC response or sensitivity to toxins and antibiotics produced by P.  
XX CC luminescens. Cells transformed to express the genes are useful for  
XX CC recombinant production of the proteins, particularly toxins and  
XX CC antibacterials useful as insecticides, bactericides and fungicides. The  
XX CC genes, proteins, vectors containing the genes and Ab are also useful  
XX CC therapeutically to treat microbial infection by bacteria or fungi that  
XX CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as  
XX CC biopesticides. Other uses of the genes and the proteins are as virulence  
XX CC factors and for identifying targets of human diseases for which P.  
XX CC luminescens is a model (particularly plague and whooping cough). This  
XX CC sequence represents one of the isolated P. luminescens proteins  
XX  
SQ Sequence 353 AA;  
Query Match 74.5%; Score 35; DB 6; Length 353;  
Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 MEQQHYV 9  
Db 331 VEQQHYV 337  
RESULT 11  
AAU15919  
ID AAU15919 standard; protein; 153 AA.  
XX  
XX AAU15919;  
XX  
XX 07-NOV-2001 (first entry)  
XX  
XX Human novel secreted protein, Seq ID 872.  
XX  
XX

KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
KW cardiac; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; anglogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
KW preservative; antiproliferative.  
XX  
OS Homo sapiens.  
XX  
PN WO200155322-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001341.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249266P.  
PR 17-NOV-2000; 2000US-0249267P.  
PR 17-NOV-2000; 2000US-0249268P.  
PR 17-NOV-2000; 2000US-0249269P.  
PR 17-NOV-2000; 2000US-0249270P.  
PR 17-NOV-2000; 2000US-0249271P.  
PR 17-NOV-2000; 2000US-0249272P.  
PR 17-NOV-2000; 2000US-0249273P.  
PR 17-NOV-2000; 2000US-0249274P.  
PR 17-NOV-2000; 2000US-0249275P.  
PR 17-NOV-2000; 2000US-0249276P.  
PR 17-NOV-2000; 2000US-0249277P.  
PR 17-NOV-2000; 2000US-0249278P.  
PR 17-NOV-2000; 2000US-0249279P.  
PR 17-NOV-2000; 2000US-0249280P.  
PR 17-NOV-2000; 2000US-0249281P.  
PR 17-NOV-2000; 2000US-0249282P.  
PR 17-NOV-2000; 2000US-0249283P.  
PR 17-NOV-2000; 2000US-0249284P.  
PR 17-NOV-2000; 2000US-0249285P.  
PR 17-NOV-2000; 2000US-0249286P.  
PR 17-NOV-2000; 2000US-0249287P.  
PR 17-NOV-2000; 2000US-0249288P.  
PR 17-NOV-2000; 2000US-0249289P.  
PR 17-NOV-2000; 2000US-0249290P.  
PR 17-NOV-2000; 2000US-0249291P.  
PR 17-NOV-2000; 2000US-0249292P.  
PR 17-NOV-2000; 2000US-0249293P.  
PR 17-NOV-2000; 2000US-0249294P.  
PR 17-NOV-2000; 2000US-0249295P.  
PR 17-NOV-2000; 2000US-0249296P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249298P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.

PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 03-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-488783/53.  
XX N-PSDB; AAS25906.  
XX New nucleic acid molecules encoding 461 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives.  
XX Claim 11; SEQ ID NO 872; 980pp; English.  
XX The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Antibodies to the proteins can also be used in  
CC alleviating symptoms associated with the disorders and in diagnostic  
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays  
CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,  
CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
CC and many other disorders listed in the specification. The polypeptides  
CC can also be used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present  
CC sequence represents a novel secreted protein of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
Query Match 72.3%; Score 34; DB 4; Length 153;  
Best Local Similarity 75.0%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 LMEQQHYV 9  
Db 113 VMAQQHYV 120  
RESULT 12  
ABU54988  
ID ABU54988 standard; protein; 153 AA.  
XX AC ABU54988;  
XX DT 18-MAR-2003 (first entry)  
XX DE Human novel polypeptide #75.  
XX Human; neural disorder; immune system disorder; renal disorder;  
KW muscular disorder; respiratory disease; reproductive disorder;  
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
KW haemostatic; antiarteriosclerotic.

XX OS Homo sapiens.  
XX US2002132753-A1.  
XX 19-SEP-2002.  
XX 17-JAN-2001; 2001US-00764864.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 22-AUG-2000; 2000US-0225758P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228927P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 03-SEP-2000; 2000US-0229509P.  
PR 03-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239355P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
XX (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.  
XX Rosen CA, Ruben SM, Barash SC;  
PI WPI; 2003-147444/14.  
XX N-PSDB; ABX73247.  
XX New polypeptides and nucleic acids, useful in gene therapy for treating,  
PT inhibiting or preventing e.g. neural, immune system, muscular,  
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
PT renal disorders.

PS Claim 11; SEQ ID NO 872; 402pp; English.  
 XX  
 CC The invention relates to human novel polypeptides and their associated  
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
 CC therapy for treating, inhibiting or preventing neural disorders, immune  
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and  
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
 CC infarction) and cancerous diseases. Sequences ABUS4914-ABUS5699 and  
 CC ABUS5748 represent human novel polypeptides of the invention  
 XX  
 SQ Sequence 153 AA;  
 Query Match 72.3%; Score 34; DB 6; Length 153;  
 Best Local Similarity 75.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LMEQOHYV 9  
 Db 113 VMAQQHYV 120  
 RESULT 13  
 ADJ68967  
 ID ADJ68967 standard; protein; 294 AA.  
 XX  
 AC ADJ68967;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Human heat mitochondrial protein as a therapeutic target SeqID773.  
 XX  
 KW mitochondrial; human; screening assay; diabetes mellitus;  
 KW Huntington's disease; osteoarthritis;  
 KW Leber's hereditary optic neuropathy; LHON;  
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
 KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;  
 KW osteopathic; ophthalmological; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003087768-A2.  
 XX  
 PD 23-OCT-2003.  
 XX  
 XX 04-APR-2003; 2003WO-US010870.  
 XX  
 XX 12-APR-2002; 2002US-0372843P.  
 PR 17-JUN-2002; 2002US-0389987P.  
 PR 20-SEP-2002; 2002US-0412418P.  
 XX  
 PA (MITO-) MITOKOR.  
 PA (BUCK-) BUCK INST AGE RES.  
 XX  
 XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;  
 PI Warnock DB;  
 XX WPI; 2003-845369/78.  
 XX  
 XX Identifying a mitochondrial target for drug screening assays and for  
 PT treating diseases associated with altered mitochondrial function,  
 PT comprises detecting a modified polypeptide in a sample and correlating  
 PT with the disease.  
 XX

PS Claim 1; SEQ ID NO 773; 180pp; English.  
 XX  
 CC This invention relates to novel mitochondrial targets that can be used  
 CC for therapeutic intervention in treating a disease associated with  
 CC altered mitochondrial function. Specifically, it refers to a method for  
 CC identifying proteins of the human heart mitochondrial proteome that are  
 CC useful for drug screening assays, as well as therapeutic targets. The  
 CC present invention describes a method for identifying such proteins that  
 CC can be used in the treatment of various diseases associated with altered  
 CC mitochondrial function including diabetes mellitus, Huntington's disease,  
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 CC compositions have neuroprotective, nontropic, antidiabetic,  
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 CC cytostatic activities. This polypeptide sequence is a human heart  
 CC mitochondrial protein of the invention.  
 XX  
 SQ Sequence 294 AA;  
 Query Match 72.3%; Score 34; DB 7; Length 294;  
 Best Local Similarity 75.0%; Pred. No. 2e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LMEQOHYV 9  
 Db 198 VMAQQHYV 205  
 RESULT 14  
 AAB41960  
 ID AAB41960 standard; protein; 302 AA.  
 XX  
 AC AAB41960;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF1724 polypeptide sequence SEQ ID NO:3448.  
 XX  
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnery; antipsoriatic; antiparkinsonian; nontropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 XX 31-MAR-2000; 2000WO-US008621.  
 PF  
 XX 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 XX Shimkets RA, Leach M;  
 PI WPI; 2000-602362/57.  
 XX  
 DR N-PSDB; AAC76169.  
 DR

```

XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 11; Page 2624; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antiparasitic; antiparkinsonian; neurotropic; neuroprotective; osteopathic;
CC cardiovulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 302 AA;

Query Match 72.3%; Score 34; DB 3; Length 302;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMEQQHYV 9
Db 206 VMAQQHYV 213
:| |||||
:| |||||

RESULT 15
ADQ67763
ID ADQ67763 standard; protein; 319 AA.
XX
AC ADQ67763;
XX
DT 07-OCT-2004 (first entry)
XX
DE Novel human protein sequence #2429.
XX
KW osteopathic; neuroprotective; neurotropic; antiparkinsonian; cytostatic;
KW gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX
OS Homo sapiens.
XX
FN EP1440981-A2.
XX
PD 28-JUL-2004.
XX
PF 21-JAN-2004; 2004EP-00001196.
XX
PR 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX
DR WPI; 2004-535376/52.
DR N-PSDB; ADQ67456.
XX

```

```

PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 4924; 2449pp; English.
XX
CC The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a protein
CC sequence of the invention.
XX
SQ Sequence 319 AA;

Query Match 72.3%; Score 34; DB 8; Length 319;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMEQQHYV 9
Db 223 VMAQQHYV 230
:| |||||
:| |||||

Search completed: November 18, 2005, 01:01:06
Job time : 54.7742 secs

```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:11:06 ; Search time 8.90323 Seconds  
(without alignments)  
97.263 Million cell updates/sec

Title: US-10-006-177-5

Perfect score: 47

Sequence: 1 ALMEQOHV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	47	100.0	769	2 A41029	integrin beta-8 ch
2	35	74.5	125	2 A69354	rubredoxin-NAD+ re
3	35	74.5	334	2 AH0300	conserved hypotet
4	34	72.3	3175	1 RRWVEV	genome polypeptid
5	33	70.2	436	2 S29891	preprotein translo
6	33	70.2	881	2 S37895	L4SI protein precu
7	33	70.2	1099	2 G90546	conserved hypotet
8	33	70.2	1998	2 T13009	hypothetical prote
9	33	70.2	2359	2 E86483	probable acetyl-Co
10	32	68.1	133	2 T44627	hypothetical prote
11	32	68.1	223	2 AE2461	hypothetical prote
12	32	68.1	365	2 E81838	hypothetical prote
13	32	68.1	365	2 G81102	hypothetical prote
14	32	68.1	481	2 G82125	Na+/H+ antiporter
15	32	68.1	519	2 B45697	immediate-early pr
16	32	68.1	565	2 A45697	neurogenesis regul
17	32	68.1	655	2 A29945	probable protein p
18	32	68.1	1428	2 T13926	polyketide synthas
19	32	68.1	2109	2 T17490	actinidin (EC 3.4
20	31	66.0	110	2 A61061	ubiquitin / riboso
21	31	66.0	155	1 UQW07A	conserved hypotet
22	31	66.0	210	2 AC0460	conserved hypotet
23	31	66.0	314	2 G90355	group 2 sigma 70-t
24	31	66.0	316	2 AC2500	thiamin-monophosph
25	31	66.0	334	2 AB2098	ferrochelatase (im
26	31	66.0	388	2 AD2274	flagellum-specific
27	31	66.0	457	2 H90963	H+-transporting tw
28	31	66.0	457	2 B64958	mannose-1-phosphat
29	31	66.0	471	2 S27676	

30 31 66.0 472 2 E47677  
31 31 66.0 496 2 H85811  
32 31 66.0 672 2 AC0422  
33 31 66.0 769 2 T49911  
34 31 66.0 783 2 T00782  
35 31 66.0 859 2 S51646  
36 31 66.0 897 2 G84613  
37 31 66.0 1029 2 F86359  
38 31 66.0 1324 2 T14070  
39 31 66.0 2415 1 A33733  
40 31 66.0 4367 1 B54802  
41 30 63.8 199 2 S65071  
42 30 63.8 205 2 T14386  
43 30 63.8 214 2 C82556  
44 30 63.8 229 2 D38941  
45 30 63.8 289 2 AH3113

#### ALIGNMENTS

##### RESULT 1

A41029

integrin beta-8 chain precursor - human

C;Species: Homo sapiens (man)

C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Aug-2004

C;Accession: A41029

R;Moyle, M.; Napier, M.A.; McLean, J.W.

J. Biol. Chem. 266, 19650-19658, 1991

A;Title: Cloning and expression of a divergent integrin subunit beta-8.

A;Reference number: A41029; MUID:92011767; PMID:1918072

A;Accession: A41029

A;Molecule type: mRNA

A;Residues: 1-769 <MOY>

A;Cross-references: UNIPROT:P26012; GB:M73780; NID:G184520; PIDN:AAA36034.1; PID:G184521

C;Superfamily: Integrin beta chain; laminin-type EGF-like homology

C;Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein

Query Match 100.0%; Score 47; DB 2; Length 769;

Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQOHV 9

Db 662 ALMEQOHV 670

##### RESULT 2

A69354

rubredoxin-NAD+ reductase (EC 1.18.1.1) - Archaeoglobus fulgidus

N;Alternate names: desulfoferrodoxin; desulfoferdoxin

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Aug-2004

C;Accession: A69354

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glocke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiaich, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: A69354

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-125 <KLE>

A;Cross-references: UNIPROT:O29425; GB:A5001047; GB:A5000782; NID:G2689370; PIDN:AAB9041

C;Comment: See PIR:RDVEX concerning activity.

C;Superfamily: Desulfoferrodoxin; desulfoferdoxin homology

C;Keywords: iron; metalloprotein; oxidoreductase

F;1-37/Domain: desulfoferdoxin homology <DSX>

F;10,13,29,30/Binding site: iron (Cys) #status predicted

Query Match 74.5%; Score 35; DB 2; Length 125;  
Best Local Similarity 71.4%; Pred. No. 5.5;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MEQQHYV 9  
|||:  
Db 71 MEQQHYI 77  
|||:

RESULT 3  
AH0300  
conserved hypothetical protein YPO2463 [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AH0300  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AH0300  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-334 <KUR>  
A;Cross-references: UNIPROT:Q8ZDU3; GB:AL590842; PIDN:G15980457; GSPDB:C  
C;Genetics: A;Gene: YPO2463

Query Match 74.5%; Score 35; DB 2; Length 334;  
Best Local Similarity 85.7%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 MEQQHYV 9  
:|||||  
Db 313 VEQQHYV 319  
:|||||

RESULT 4  
RRWVEV  
genome polyprotein - equine arteritis virus  
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C;Species: equine arteritis virus  
A;Note: host Equus caballus (domestic horse)  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 09-Jul-2004  
C;Accession: A39925; S10158; B39925  
R;Den Boon, J.A.; Snijder, E.J.; Chirnside, E.D.; De Vries, A.A.F.; Horzinek, M.C.; Spaa J. Virol. 65, 2910-2920, 1991  
A;Title: Equine arteritis virus is not a togavirus but belongs to the coronaviruslike subfamily  
A;Reference number: A39925; MUID:91237805; PMID:1851863  
A;Accession: A39925  
A;Molecule type: Genomic RNA  
A;Residues: 1-3175 <DEN>  
A;Cross-references: UNIPROT:P19811; EMBL:X53459  
A;Note: a -1 ribosomal frameshift occurs between the codons AAC for 1727-Aen and CUG for Rde Vries, A.A.F.; Chirnside, E.D.; Bredendbeek, P.V.; Gravestien, L.A.; Horzinek, M.C.; Nucleic Acids Res. 18, 3241-3247, 1990  
A;Title: All subgenomic mRNAs of equine arteritis virus contain a common leader sequence  
A;Reference number: S10158; MUID:90287699; PMID:2162519  
A;Accession: S10158  
A;Status: translation not shown  
A;Molecule type: Genomic RNA  
A;Residues: 1-17 <VRI>  
A;Cross-references: EMBL:X52277  
C;Superfamily: equine arteritis virus RNA-directed RNA polymerase  
C;Keywords: nucleotidyltransferase

Query Match 72.3%; Score 34; DB 1; Length 3175;  
Best Local Similarity 62.5%; Pred. No. 3.3e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALMEQQHY 8  
|||:|:|

Db 587 ALLEEYH 594

RESULT 5  
S29891  
preprotein translocase secY - Micrococcus luteus  
C;Species: Micrococcus luteus, Micrococcus lysodeikticus  
C;Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jul-2004  
C;Accession: S29891  
R;Ohama, T.; Muto, A.; Osawa, S.  
J. Mol. Evol. 29, 381-395, 1989  
A;Title: Spectinomycin operon of Micrococcus luteus: evolutionary implications of organellar  
A;Reference number: S17070; MUID:90133967; PMID:2533272  
A;Accession: S29891  
A;Molecule type: DNA  
A;Residues: 1-436 <OHA>  
A;Cross-references: UNIPROT:P33108; EMBL:X17524; NID:G44425; PIDN:CAA35567.1; PID:G44433  
C;Genetics: A;Gene: secY  
A;Start codon: GTG  
C;Superfamily: preprotein translocase secY  
C;Keywords: protein transport; transmembrane protein  
F;18-39/Domain: transmembrane #status predicted <TM01>  
F;72-92/Domain: transmembrane #status predicted <TM02>  
F;119-136/Domain: transmembrane #status predicted <TM03>  
F;157-175/Domain: transmembrane #status predicted <TM04>  
F;182-204/Domain: transmembrane #status predicted <TM05>  
F;214-232/Domain: transmembrane #status predicted <TM06>  
F;266-286/Domain: transmembrane #status predicted <TM07>  
F;315-332/Domain: transmembrane #status predicted <TM08>  
F;372-395/Domain: transmembrane #status predicted <TM09>

Query Match 70.2%; Score 33; DB 2; Length 436;  
Best Local Similarity 75.0%; Pred. No. 58;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALMEQQHY 8  
|:|:|:  
Db 424 AQMEQRHY 431  
|:|:|:

RESULT 6  
S37895  
LHSL protein precursor - Yeast (Saccharomyces cerevisiae)  
N;Alternate names: probable endoplasmic reticulum protein; protein YKL073w; protein YKL3  
C;Species: Saccharomyces cerevisiae  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
C;Accession: S37895; S37898; S39169; S44514  
R;Rasmussen, S.; von Wettstein, D.  
submitted to the Protein Sequence Database, March 1994  
A;Reference number: S37872  
A;Accession: S37895  
A;Molecule type: DNA  
A;Residues: 1-881 <RAS>  
A;Cross-references: UNIPROT:P36016; EMBL:Z28073; NID:G486102; MIPS:YKL073w  
A;Experimental source: strain S288C  
R;Pohl, T.M.; Pohl, F.M.  
submitted to the Protein Sequence Database, March 1994  
A;Reference number: S37897  
A;Accession: S37898  
A;Molecule type: DNA  
A;Residues: 1-881 <POH>  
A;Cross-references: EMBL:Z28073; NID:G486102; PID:G486103; MIPS:YKL073w  
A;Experimental source: strain S288C  
R;Rasmussen, S.W.  
submitted to the EMBL Data Library, November 1993  
A;Reference number: S39168  
A;Accession: S39169  
A;Molecule type: DNA  
A;Residues: 1-881 <RA2>  
A;Cross-references: EMBL:X75780; NID:G433625; PID:G433627  
R;Rasmussen, S.W.  
Yeast 10, 69-74, 1994



A:Title: Sequence of a 20.7 kb region of yeast chromosome XI includes the NUP100 gene, a xp in addition to seven ORFs with weak or no significant similarity to known proteins.  
 A:Reference number: S44513

A:Accession: S44514

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-881 <RA3>

A:CROSS-references: EMBL:X75780; NID:g433625; PID:g433627

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993

C:Genetics:

A:Gene: SGD:LHS1; CER1; SS11

A:CROSS-references: SGD:S0001556; MIPS:YKL073w

A:Map position: 11L

C:Keywords: endoplasmic reticulum

F:1-20/Domain: signal sequence #status predicted <STG>

F:878-881/Region: endoplasmic reticulum retention signal

Query Match 70.2%; Score 33; DB 2; Length 881;  
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EQQHVV 9

Db 229 EQQHVI 234

RESULT 7

G90546 conserved hypothetical protein MYPU\_2790 [imported] - Mycoplasma pulmonis (strain UAB CT)

C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C:Accession: G90546

R:Chambaud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm

A:Reference number: A95512; MUID:21267165; PMID:11353084

A:Accession: G90546

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1099 <KUR>

A:CROSS-references: UNIPROT:Q98QT2; GB:AL445566; PID:g14089692; PIDN:CAC13452.1; GSPDB:G

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPU 2790

A:Genetic code: SGC3

Query Match 70.2%; Score 33; DB 2; Length 1099;  
 Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQQH 7

Db 537 ALLEQQH 543

RESULT 8

T13009 hypothetical protein T24C20.80 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004

C:Accession: T13009

R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa

submitted to the Protein Sequence Database, July 1999

A:Reference number: Z17586

A:Accession: T13009

A:Molecule type: DNA

A:Residues: 1-1998 <CHO>

A:CROSS-references: UNIPROT:Q9STR8; EMBL:AL096856; GSPDB:GN00061; ATSP:T24C20.80

A:Experimental source: cultivar Columbia; BAC clone T24C20

C:Genetics:

A:Gene: ATSP:T24C20.80

A:Map position: 3

A:Introns: 71/3; 117/1; 162/3; 230/1; 309/3; 348/2; 396/1; 452/3; 532/2; 604/1; 1005/2;

C:Superfamily: Arabidopsis thaliana hypothetical protein T24C20.80

Query Match 70.2%; Score 33; DB 2; Length 1998;  
 Best Local Similarity 62.5%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMEQQHVV 9

Db 1706 LLEAQHYI 1713

RESULT 9

E86483

probable acetyl-CoA carboxylase, 9984-22276 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: E86483

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

ansen, N.F.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,

Chin, C.W.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

A:; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86483

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2359 <STO>

A:CROSS-references: UNIPROT:Q9C8G0; GB:AE005172; NID:g10645483; PIDN:AAG21596.1; GSPDB:G

C:Genetics:

A:Map position: 1

C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-

Query Match 70.2%; Score 33; DB 2; Length 2359;  
 Best Local Similarity 75.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMEQQHVV 9

Db 1351 LMEQYVV 1358

RESULT 10

T44627

hypothetical protein [imported] - Butyrivibrio fibrisolvens (fragment)

C:Species: Butyrivibrio fibrisolvens

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T44627

R:Dalrymple, B.P.; Swadling, Y.; Cybinski, D.H.; Xue, G.P.

FEMS Microbiol. Lett. 143, 115-120, 1996

A:Title: Cloning of a gene encoding cinnamoyl ester hydrolase from the ruminal bacterium

A:Reference number: Z22816; MUID:96434496; PMID:8837463

A:Accession: T44627

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-133 <DAL>

A:CROSS-references: UNIPROT:P70887; EMBL:U44893; NID:g1622730; PIDN:AAC44496.1; PID:g162

A:Experimental source: strain E14

Query Match 68.1%; Score 32; DB 2; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMEQQH 7

Db 4 LMEQQH 9

RESULT 11



Db       ||:|:|:|  
         82 LNDEQHFI 89

Search completed: November 18, 2005, 01:09:32  
Job time : 12.0032 secs

**This Page Blank (user)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:04:26 ; Search time 39.9677 Seconds  
(without alignments)  
115.311 Million cell updates/sec

Title: US-10-006-177-5

Perfect score: 47

Sequence: 1 ALMEQQHV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	100.0	769	1 ITB8_HUMAN	P26012 homo sapien
2	38	80.9	565	2 Q9GV73	Q9GV73 halocynthia
3	36	76.6	930	2 Q62L49	Q62L49 burkholderi
4	36	76.6	930	2 Q63T87	Q63T87 burkholderi
5	35	74.5	125	1 DESR_ARCPU	O29425 archaeoglob
6	35	74.5	232	2 Q7UW9	Q7UW9 rhodospirill
7	35	74.5	334	2 Q669H9	Q669H9 yersinia ps
8	35	74.5	334	2 Q8ZDU3	Q8ZDU3 yersinia pe
9	35	74.5	335	2 Q9CNW6	Q9CNW6 pasteurella
10	35	74.5	337	2 Q7N5Y4	Q7N5Y4 photorhabd
11	35	74.5	476	2 Q8XSC7	Q8XSC7 ralstonia s
12	35	74.5	562	2 Q80YT3	Q80YT3 mus musculu
13	34	72.3	291	2 Q8MDA2	Q8MDA2 paramachlyd
14	34	72.3	294	2 Q88CV9	Q88CV9 homo sapien
15	34	72.3	294	2 Q9UL40	Q9UL40 homo sapien
16	34	72.3	319	2 Q6ZMW1	Q6ZMW1 homo sapien
17	34	72.3	337	2 Q6D0W6	Q6D0W6 erwinia car
18	34	72.3	366	2 Q6FV33	Q6FV33 candida gla
19	34	72.3	417	2 Q6F7C4	Q6F7C4 acinetobact
20	34	72.3	3175	1 RFOA_EAV	P19811 equine arte
21	33	70.2	86	1 Y043_OCEIH	Q9EU46 oceanobacil
22	33	70.2	160	2 Q6AXC5	Q6AXC5 mus musculu
23	33	70.2	198	2 Q48371	Q48371 insertion s
24	33	70.2	229	2 Q9H898	Q9H898 homo sapien
25	33	70.2	229	2 Q8BZ94	Q8BZ94 mus musculu
26	33	70.2	289	2 Q75F51	Q75F51 ashbya goss
27	33	70.2	301	2 Q8T8X3	Q8T8X3 drosophila
28	33	70.2	302	2 Q8T8X3	Q8T8X3 drosophila
29	33	70.2	302	2 Q9VRB5	Q9VRB5 drosophila
30	33	70.2	354	2 Q83D88	Q83D88 coxilla bu
31	33	70.2	366	2 Q9F886	Q9F886 wautersia m

32	33	70.2	393	1 NUCC_PSINU	Q8WHX3 psilotum nu
33	33	70.2	435	2 Q75DK5	Q75DK5 ashbya goss
34	33	70.2	436	1 SECY_MICLU	P33108 micrococcus
35	33	70.2	471	2 Q6FFD3	Q6FFD3 acinetobact
36	33	70.2	526	2 Q6E5A8	Q6E5A8 suberites d
37	33	70.2	619	2 Q6ZMW9	Q6ZMW9 homo sapien
38	33	70.2	724	2 Q9GTU8	Q9GTU8 hydra atten
39	33	70.2	775	2 Q8RW99	Q8RW99 arabidopsis
40	33	70.2	881	1 LHS1_YEAST	P36016 saccharomyc
41	33	70.2	885	2 Q859H5	Q859H5 pseudomonas
42	33	70.2	902	2 Q7PMF0	Q7PMF0 anopheles g
43	33	70.2	938	2 Q8RWV7	Q8RWV7 arabidopsis
44	33	70.2	1099	2 Q88QT2	Q88QT2 mycoplasma
45	33	70.2	1318	2 Q6ZUA9	Q6ZUA9 homo sapien

#### ALIGNMENTS

##### RESULT 1

ITB8\_HUMAN  
ID ITB8\_HUMAN STANDARD; PRT; 769 AA.  
AC P26012;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Integrin beta-8 precursor.  
GN Name=ITGB8;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=92011767; PubMed=1918072;  
RA Moyle M., Napier M.A., McLean J.W.;  
RT "Cloning and expression of a divergent integrin subunit beta 8.";  
RL J. Biol. Chem. 266:19650-19658(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Tin-Mollam A., Sutterer C., Fronick B.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Integrin alpha-V/beta-8 is a receptor for fibronectin.  
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Beta-8 associates with alpha-V.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Placenta, kidney, brain, ovary, uterus and in several transformed cells. Transiently expressed in 293 human embryonic kidney cells.  
CC -!- SIMILARITY: Belongs to the integrin beta chain family.  
CC -!- SIMILARITY: Contains 1 VWFA-like domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
-----

EMBL; M73780; AAA36034.1; --  
EMBL; AC004130; AAQ96845.1; --  
PIR; A41029; 1L5G.  
HSSP; P05106; 1L5G.  
Genew; HGNC:6163; ITGB8.  
MIM; 604160; --  
GO; GO:0008305; C:integrin complex; TAS.  
GO; GO:0007155; P:cell adhesion; TAS.  
InterPro; IPR006209; EGF-like.  
InterPro; IPR002369; Integrin\_B.  
InterPro; IPR001169; Integrin\_beta\_C.  
Pfam; PF00362; Integrin\_beta; 1.

```
DR PRINTS; PR01186; INTEGRINB.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN 1.
DR PROSITE; PS00243; INTEGRIN BETA; 2.
KW Cell adhesion; Glycoprotein; Integrin; Receptor; Repeat; Signal;
KW Transmembrane.
FT SIGNAL 1 42 Potential.
FT CHAIN 43 769 Integrin beta-8.
FT DOMAIN 43 684 Extracellular (Potential).
FT TRANSMEM 685 704 Potential.
FT DOMAIN 705 769 Cytoplasmic (Potential).
FT DOMAIN 146 384 WWFA-like.
FT DOMAIN 471 629 4 cysteine-rich tandem repeats.
FT REPEAT 471 510 I.
FT REPEAT 511 552 II.
FT REPEAT 553 592 III.
FT REPEAT 593 629 IV.
FT DISULFID 47 469 By similarity.
FT DISULFID 55 65 By similarity.
FT DISULFID 58 94 By similarity.
FT DISULFID 68 83 By similarity.
FT DISULFID 211 218 By similarity.
FT DISULFID 266 307 By similarity.
FT DISULFID 407 419 By similarity.
FT DISULFID 467 471 By similarity.
FT DISULFID 494 499 By similarity.
FT DISULFID 526 531 By similarity.
FT DISULFID 528 561 By similarity.
FT DISULFID 533 546 By similarity.
FT DISULFID 567 572 By similarity.
FT DISULFID 574 583 By similarity.
FT DISULFID 585 593 By similarity.
FT DISULFID 607 612 By similarity.
FT DISULFID 609 657 By similarity.
FT DISULFID 614 624 By similarity.
FT DISULFID 627 630 By similarity.
FT DISULFID 634 643 By similarity.
FT CARBOHYD 233 233 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 402 402 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 421 421 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 431 431 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 456 456 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 466 466 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 648 648 N-linked (GlcNAc. .) (Potential).
SQ SEQUENCE 769 AA; 85631 MW; F7E3994F92B12A65 CRC64;

Query Match 100.0%; Score 47; DB 1; Length 769;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQOHYV 9
Db 662 ALMEQOHYV 670

RESULT 2
Q9GV73 PRELIMINARY; PRT; 565 AA.
ID Q9GV73
AC Q9GV73;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE HNF-3.
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyuridae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gastrula;
RX MEDLINE=98145942; PubMed=9486537; DOI=10.1016/S0925-4773(97)00162-7;
RA Shimauchi Y., Yasuo H., Satoh N.;
RT "Autonomy of ascidian fork head/HNF-3 gene expression.";
```

```
RL Mech. Dev. 69:143-154(1997).
DR EMBL; AB007406; BAB16309.1; -.
DR HSP; Q99598; ID5V.
DR GO; GO:0005834; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1..DNA_bnd.
DR PRINTS; PR00053; FORKHEAD.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS00039; FORK_HEAD_3; 1.
SQ SEQUENCE 565 AA; 62818 MW; 1BBACB93371EEAB CRC64;

Query Match 80.9%; Score 38; DB 2; Length 565;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQOHYV 9
Db 315 ALLEQHHI 323

RESULT 3
Q62L49 PRELIMINARY; PRT; 930 AA.
ID Q62L49
AC Q62L49;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Maltooligosyl trehalose synthase, putative.
GN ORFNames=BMA0816;
OS Burkholderia mallei ATCC 23344.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=243160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23344;
RA Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C.,
RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,
RA Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,
RA Mohammed Y., Nelson W.C., Radune D., Romero C.M., Sarria S.,
RA Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y., Zafar N.,
RA Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL; CP000010; AAU49513.1; -.
SQ SEQUENCE 930 AA; 100833 MW; FDBB168AAE710793 CRC64;

Query Match 76.6%; Score 36; DB 2; Length 930;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMEQOHY 8
Db 250 LLEQOHY 256

RESULT 4
Q63T87 PRELIMINARY; PRT; 930 AA.
ID Q63T87
AC Q63T87;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative glycosyl hydrolase.
GN ORFNames=BPSL2080;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

-----

CC EMBL; A5001047; AAB90417.1; -.  
 CC FIR; A69354; A69354.  
 CC HSSP; P22076; 1DFX.  
 CC TIGR; AF0833; -.  
 CC InterPro; IPR004462; Desulf\_FeS4.  
 CC InterPro; IPR002742; Desulf\_FeS4.  
 CC InterPro; IPR004793; Dfx\_rbo.  
 CC Pfam; PF06397; Desulfoferrodox\_N; 1.  
 CC Pfam; PF01880; Desulfoferrodox; 1.  
 CC ProDom; PD006618; Desulfoferrodox; 1.  
 CC TIGRfam; TIGR00319; desulf\_FeS4; 1.  
 CC TIGRfam; TIGR00320; dfx\_rbo; 1.  
 CC TIGRfam; TIGR00332; neela\_ferrous; 1.  
 CC Complete proteome; Electron transport; Hypothetical protein; Iron.  
 CC METAL 10 10 Iron 1 (By similarity).  
 CC METAL 13 13 Iron 1 (By similarity).  
 CC METAL 29 29 Iron 1 (By similarity).  
 CC METAL 30 30 Iron 1 (By similarity).  
 CC METAL 49 49 Iron 2 (By similarity).  
 CC METAL 69 69 Iron 2 (By similarity).  
 CC METAL 75 75 Iron 2 (By similarity).  
 CC METAL 116 116 Iron 2 (By similarity).  
 CC METAL 119 119 Iron 2 (By similarity).  
 CC SEQUENCE 125 AA; 14115 MW; 75C815440A871661 CRC64;  
 Query Match 74.5%; Score 35; DB 1; Length 125;  
 Best Local Similarity 71.4%; Pred. No. 30;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MEQOHV 9  
 DB 71 MEEQHYI 77

RESULT 6  
 Q7UWN9 PRELIMINARY; PRT; 232 AA.  
 AC Q7UWN9;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=RB1897;  
 OS Rhodopirellula baltica.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OC NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=2275913; PubMed=12835416; DOI=10.1073/pnas.1431443100;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
 RA Schlesner H., Amann K., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 RT strain 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
 DR EMBL; BX294136; CAD72323.1; -.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 232 AA; 25778 MW; F8CC4171B8CD60DC CRC64;  
 Query Match 74.5%; Score 35; DB 2; Length 232;  
 Best Local Similarity 75.0%; Pred. No. 59;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALMEQHY 8  
 DB 110 ALLEQSHY 117

Burkholderiaceae; Burkholderia.  
 NCBI\_TaxID=272560;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=K96243;  
 PubMed=15377794;  
 Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.M.,  
 Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,  
 Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,  
 Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,  
 Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,  
 Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,  
 Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,  
 Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,  
 Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,  
 Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;  
 "Genomic plasticity of the causative agent of melioidosis,  
 Burkholderia pseudomallei.";  
 Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).  
 RL EMBL; BX571965; CAH36081.1; -.  
 KW Hydrolase.  
 SQ SEQUENCE 930 AA; 100866 MW; F9C8E5E3570DD977 CRC64;  
 Query Match 76.6%; Score 36; DB 2; Length 930;  
 Best Local Similarity 85.7%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMEQOHY 8  
 DB 250 LLEQOHY 256

RESULT 5  
 DESR\_ARCFU STANDARD; PRT; 125 AA.  
 ID DESR\_ARCFU  
 AC Q29425;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Desulfoferrodoxin homolog (Dfx).  
 GN OrderedLocusNames=AF0833;  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 OC NCBI\_TaxID=2234;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=9389475; DOI=10.1038/37052;  
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kervilange A.R., Graham D.E., Kyrpides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,  
 RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A.,  
 RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,  
 RA Utterback T.R., Cotton M.D., Spriggs T., Artiach P., Kaine B.P.,  
 RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,  
 RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,  
 RA Woese C.R., Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 RT reducing archaeon Archaeoglobus fulgidus.";  
 RL Nature 390:364-370(1997)  
 CC -1- FUNCTION: Non-heme iron protein, contains two types of iron atoms  
 CC per molecule: a desulfoferrodoxin-like FeS(4) site and an  
 CC octahedral coordinated high-spin ferrous site with nitrogen/  
 CC oxygen-containing ligands (By similarity).  
 CC -1- SIMILARITY: Belongs to the desulfoferrodoxin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

```
RESULT 7
Q669H9 ID Q669H9 PRELIMINARY; PRT; 334 AA.
AC Q669H9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=YPTB2506;
OS Yersinia pseudotuberculosis IP 32953.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=273123;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP 32953;
RX PubMed=15358858;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin L.V.,
RA Brubaker R.R., Fowler J., Hinnebusch B.J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the genome evolution of Yersinia pestis through whole
RT genome comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
RW EMBL; EX936398; CAH21744.1; -.
KW Hypothetical protein.
SQ SEQUENCE 334 AA; 36930 MW; EACC853665F6235F CRC64;

Query Match 74.5%; Score 35; DB 2; Length 334;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MEQQHYV 9
Db 313 VEQQHYV 319

RESULT 8
Q8ZDU3 ID Q8ZDU3 PRELIMINARY; PRT; 334 AA.
AC Q8ZDU3; Q74TA7; Q7CJ38;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein YPO2463.
GN OrderedLocusNames=YP02282, YPO2463, Y1726;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis; DOI=10.1038/35097083;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
```

```
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ414152; CAC91268.1; -.
DR EMBL; AE013776; AAM85294.1; -.
DR EMBL; AE017135; AAS62488.1; -.
DR PIR; AH0300; AH0300.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 334 AA; 36901 MW; 7D729539A3C1235B CRC64;

Query Match 74.5%; Score 35; DB 2; Length 334;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MEQQHYV 9
Db 313 VEQQHYV 319

RESULT 9
Q9CNW6 ID Q9CNW6 PRELIMINARY; PRT; 335 AA.
AC Q9CNW6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein PM0307.
GN OrderedLocusNames=PM0307;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006066; AAK02391.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 335 AA; 37843 MW; 7AA6BFF6400A4B70 CRC64;

Query Match 74.5%; Score 35; DB 2; Length 335;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MEQQHYV 9
Db 313 VEQQHYV 319

RESULT 10
Q7NSY4 ID Q7NSY4 PRELIMINARY; PRT; 337 AA.
AC Q7NSY4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to unknown protein.
GN OrderedLocusNames=plu1795;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
```



```

OX NCBI_TaxID=141679;
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Ruenik K., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Medigue C., Lanols A., Powell K., Siguiet P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photobacterium
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313 (2003).
DR EMBL: BX571865; CAE14088.1; -.
DR Photobact; plui795; -.
KW Complete proteome.
SQ SEQUENCE 337 AA; 37361 MW; 63985886D185029 CRC64;

Query Match 74.5%; Score 35; DB 2; Length 337;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MEQQHYV 9
Db 316 VEQQHYV 322
:|||||
:|||||

RESULT 11
Q8XSC7 PRELIMINARY; PRT; 476 AA.
AC Q8XSC7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PROBABLE HYDROLASE TRANSMEMBRANE PROTEIN (EC 3.-.-.-).
GN Name=RS03928; OrderedLocusNames=RS0551;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguiet P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502 (2002).
DR EMBL: AL646079; CAD17702.1; -.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0016787; F: hydrolase activity; IEA.
DR Pfam: PF01979; Amidohydro 1; 1.
DR ProDom: PD001248; Amidohydro like; 1.
KW Complete proteome; Hydrolase; Plasmid; Transmembrane.
SQ SEQUENCE 476 AA; 51227 MW; 7F3ECCBDFCC5C75F CRC64;

Query Match 74.5%; Score 35; DB 2; Length 476;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMEQQHYV 9
Db 266 LMNQQHYL 273
|||||
|||||

RESULT 12
Q80YT3 PRELIMINARY; PRT; 562 AA.
ID Q80YT3

```

```

AC Q80YT3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gene model 443.
CN Name=Gm443;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schin J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL: BC050797; AAH50797.1; -.
DR MGD: MGI:2685289; Gm43.
DR GO: GO:0005524; F: ATP binding; IEA.
DR GO: GO:0051082; F: unfolded protein binding; IEA.
DR GO: GO:0006457; F: protein folding; IEA.
DR InterPro: IPR002194; Chaperonin TCP-1.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR InterPro: IPR008950; GroEL-ATPase.
DR Pfam: PF00118; Cpn60_TCP1; 1.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00750; TCP1_1; UNKNOWN_1.
KW ATP-binding; Chaperone.
SQ SEQUENCE 562 AA; 60719 MW; 9DE60624F90C3384 CRC64;

Query Match 74.5%; Score 35; DB 2; Length 562;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALMEQQHYV 9
Db 123 ALLEQAHYL 131
|||||
|||||

RESULT 13
Q6MD42 PRELIMINARY; PRT; 291 AA.
ID Q6MD42
AC Q6MD42;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Probable XerD protein.
GN Name=XerD; OrderedLocusNames=pc0783;
OS Parachlamydia sp. (strain UWE25) (subsp. Acanthamoeba sp.).
OC Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.

```

```

OX NCBI_TaxID=264201;
RN [1]
RP SEQUENCE FROM N.A.
RA Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,
RA Fattmann B., Brandt P., Nyakatura G.J., Broege M., Frishman D.,
RA Rattei T., Mewes H.-W., Wagner M.;
RT "Genome sequence of an amoeba symbiont and its use for reconstructing
RL the evolutionary history of chlamydiae.";
DR Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX908798; CAF23507.1; -.
DR HSSP; P21891; 1AOP.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR011010; DNA_drk_join_enz.
DR InterPro; IPR010998; L_intgrase_like_N.
DR InterPro; IPR002104; Phage integrase.
DR InterPro; IPR004107; Phage integr N.
DR Pfam; PF00589; Phage integrase; 1.
DR Pfam; PF02899; Phage_integr_N; 1.
DR Complete proteome.
KW SEQUENCE 291 AA; 33541 MW; 22C0BE793FE6B662 CRC64;

Query Match 72.3%; Score 34; DB 2; Length 291;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALMEQQHY 8
Db 58 ALKQQHY 65

RESULT 14
ID Q68CV9 PRELIMINARY; PRT; 294 AA.
AC Q68CV9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein DKFZp547M223.
GN Name=DKFZp547M223;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA The German cDNA Consortium;
RG Tissue=Brain;
RA Bloeker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749692; CAH18475.1; -.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR003604; Znf_U1.
DR SMART; SM00355; Znf_C2H2; 3.
DR SMART; SM00451; Znf_U1; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 294 AA; 32958 MW; 543372E183D08299 CRC64;

Query Match 72.3%; Score 34; DB 2; Length 294;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMEQQHYV 9
Db 198 VMAQQHYV 205

RESULT 15
Q9UL40 PRELIMINARY; PRT; 294 AA.
ID Q9UL40

```

```

AC Q9UL40;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Double-stranded RNA-binding zinc finger protein JAZ (Zinc finger
DE protein 346).
GN Name=ZNF346;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=99419013; PubMed=10488071; DOI=10.1074/jbc.274.39.27399;
RA Yang M., May W.S., Ito T.;
RT "JAZ requires the double-stranded RNA-binding zinc finger motifs for
RT nuclear localization.";
RL J. Biol. Chem. 274:27399-27406(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083340; AAD52018.1; -.
DR EMBL; BC007775; AAH07775.1; -.
DR GO; GO:0005730; C:nucleolus; TAS.
DR GO; GO:0003725; F:double-stranded RNA binding; TAS.
DR GO; GO:0006915; P:apoptosis; TAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; TAS.
DR GO; GO:0006396; P:RNA processing; TAS.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR003604; Znf_U1.
DR SMART; SM00355; Znf_C2H2; 4.
DR SMART; SM00451; Znf_U1; 4.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_4.
SQ SEQUENCE 294 AA; 32332 MW; 095F02E19FBCF6E0 CRC64;

Query Match 72.3%; Score 34; DB 2; Length 294;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMEQQHYV 9
Db 198 VMAQQHYV 205

```

Search completed: November 18, 2005, 01:08:04  
Job time : 44.9677 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 00:48:52 : Search time 13.6452 Seconds  
(without alignments)  
49.237 Million cell updates/sec

Title: US-10-006-177-5  
Perfect score: 47

Sequence: 1 ALMEQQHYV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	769	1	US-08-454-455-6
2	47	100.0	831	4	Sequence 6, Appli
3	34	72.3	164	4	Sequence 10169, A
4	33	70.2	378	4	Sequence 14156, A
5	32	68.1	152	4	Sequence 26019, A
6	32	68.1	191	4	Sequence 7505, Ap
7	32	68.1	191	4	Sequence 12606, A
8	32	68.1	591	4	Sequence 18, Appl
9	31	66.0	79	4	Sequence 348, App
10	31	66.0	164	4	Sequence 8198, Ap
11	31	66.0	235	4	Sequence 13178, A
12	31	66.0	393	4	Sequence 13178, A
13	31	66.0	478	4	Sequence 19806, A
14	31	66.0	538	4	Sequence 7205, Ap
15	31	66.0	1208	3	Sequence 24079, A
16	31	66.0	1208	4	Sequence 2, Appli
17	30	63.8	60	4	Sequence 6481, Ap
18	30	63.8	120	4	Sequence 40935, A
19	30	63.8	120	4	Sequence 56151, A
20	30	63.8	217	4	Sequence 767-33049
21	30	63.8	256	4	Sequence 14410, A
22	30	63.8	260	4	Sequence 16370, A
23	30	63.8	294	4	Sequence 12165, A
24	30	63.8	344	4	Sequence 4890, Ap
25	30	63.8	353	4	Sequence 2888, Ap
26	30	63.8	364	4	Sequence 16427, A
27	30	63.8	402	4	Sequence 17911, A

28 30 63.8 492 4 US-09-489-039A-9018 Sequence 9018, Ap  
29 30 63.8 526 4 US-09-949-016-10277 Sequence 10277, A  
30 30 63.8 559 4 US-09-949-016-9995 Sequence 9995, Ap  
31 30 63.8 564 3 US-09-134-001C-2897 Sequence 2897, Ap  
32 30 63.8 592 4 US-09-252-991A-31642 Sequence 31642, A  
33 30 63.8 746 4 US-09-248-796A-19979 Sequence 19979, A  
34 30 63.8 881 4 US-09-252-991A-31702 Sequence 31702, A  
35 30 63.8 888 4 US-09-252-991A-25825 Sequence 25825, A  
36 30 63.8 928 4 US-09-134-000C-5307 Sequence 5307, Ap  
37 30 63.8 1150 4 US-09-252-991A-24671 Sequence 24671, A  
38 30 63.8 1154 4 US-09-489-039A-7724 Sequence 7724, Ap  
39 30 63.8 1159 4 US-09-328-352-7624 Sequence 7624, Ap  
40 30 63.8 1185 4 US-09-252-991A-18328 Sequence 18328, A  
41 30 63.8 1244 4 US-09-543-681A-6274 Sequence 6274, Ap  
42 29 61.7 40 2 US-08-248-839C-168 Sequence 168, App  
43 29 61.7 53 3 US-08-630-916A-70 Sequence 70, Appl  
44 29 61.7 64 4 US-09-621-976-4584 Sequence 4584, Ap  
45 29 61.7 135 3 US-09-091-117-1 Sequence 1, Appli

#### ALIGNMENTS

RESULT 1  
US-08-454-455-6  
; Sequence 6, Application US/08454455  
; Patent No. 5635601  
; GENERAL INFORMATION:  
; APPLICANT: Moyle, Matthew  
; APPLICANT: McLean, John W.  
; TITLE OF INVENTION: NOVEL BETA INTEGRIN SUBUNIT  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/454,455  
; FILING DATE: 30-May-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/193989  
; FILING DATE: 09-FEB-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/004142  
; FILING DATE: 13-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/670607  
; FILING DATE: 14-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 00,000  
; REFERENCE/DOCKET NUMBER: P0699C2D2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 769 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-454-455-6

Query Match

100.0%; Score 47; DB 1; Length 769;

```
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQOHYV 9
Db 662 ALMEQOHYV 670
|||||

RESULT 2
US-09-949-016-10169
; Sequence 10169, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10169
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10169

Query Match 100.0%; Score 47; DB 4; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQOHYV 9
Db 724 ALMEQOHYV 732
|||||

RESULT 3
US-09-902-540-14156
; Sequence 14156, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wisegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,893
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14156
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14156

Query Match 72.3%; Score 34; DB 4; Length 164;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMEQOHYV 9
Db 17 LMTQOHL 24
|||||
```

```
RESULT 4
US-09-248-796A-26019
; Sequence 26019, Application US/09248796A
; Patent No. 8747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26019
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-26019

Query Match 70.2%; Score 33; DB 4; Length 378;
Best Local Similarity 75.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LMEQOHYV 9
Db 131 LQEQNHV 138
|||||

RESULT 5
US-09-543-681A-7505
; Sequence 7505, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7505
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7505

Query Match 68.1%; Score 32; DB 4; Length 152;
Best Local Similarity 71.4%; Pred. No. 51;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMEQOHY 8
Db 12 LIEQEHY 18
|:|:|

RESULT 6
US-09-489-039A-12606
; Sequence 12606, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
```

; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 12606  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12606

Query Match 68.1%; Score 32; DB 4; Length 191;  
Best Local Similarity 62.5%; Pred. No. 65;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQQHY 8  
Db 8 SLMQQQHH 15  
:|:|:|:

RESULT 7  
US-09-561-077C-18  
; Sequence 18, Application US/09561077C  
; Patent No. 6706501  
; GENERAL INFORMATION:  
; APPLICANT: Rosson, Reinhardt D.  
; APPLICANT: Deng, Ming-de  
; APPLICANT: Grund, Alan D.  
; TITLE OF INVENTION: LINOLEATE ISOMERASE  
; FILE REFERENCE: 3161-20-Cl  
; CURRENT APPLICATION NUMBER: US/09/561,077C  
; CURRENT FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/141,798  
; PRIOR FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 591  
; TYPE: PRT  
; ORGANISM: Lactobacillus reuteri  
US-09-561-077C-18

Query Match 68.1%; Score 32; DB 4; Length 591;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMEQQH 7  
Db 584 LMEQQH 589  
|||||

RESULT 8  
US-09-221-014-18  
; Sequence 18, Application US/09221014C  
; Patent No. 6743609  
; GENERAL INFORMATION:  
; APPLICANT: Rosson, Reinhardt D.  
; APPLICANT: Grund, Alan D.  
; APPLICANT: Deng, Ming-de  
; APPLICANT: Sanchez-Riera, Fernando  
; TITLE OF INVENTION: LINOLEATE ISOMERASE  
; FILE REFERENCE: 3161-20  
; CURRENT APPLICATION NUMBER: US/09/221,014C  
; CURRENT FILING DATE: 1998-12-23  
; EARLIER APPLICATION NUMBER: 60/068,617  
; EARLIER FILING DATE: 1997-12-23  
; EARLIER APPLICATION NUMBER: 60/089,560  
; EARLIER FILING DATE: 1998-06-17  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 591  
; TYPE: PRT  
; ORGANISM: Lactobacillus reuteri  
US-09-221-014-18

Query Match 68.1%; Score 32; DB 4; Length 591;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMEQQH 7  
Db 584 LMEQQH 589  
|||||

RESULT 9  
US-09-732-210-348  
; Sequence 348, Application US/09732210  
; Patent No. 6573361  
; GENERAL INFORMATION:  
; APPLICANT: Bunkers, Greg J.  
; APPLICANT: Liang, Jihong  
; APPLICANT: Mittanck, Cindy A.  
; APPLICANT: Seale, Jeffrey W.  
; APPLICANT: Wu, Yonnie S.  
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
; FILE REFERENCE: 38-21(15036)B  
; CURRENT APPLICATION NUMBER: US/09/732,210  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,513  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,340  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 1753  
; SEQ ID NO 348  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Manduca sexta  
US-09-732-210-348

Query Match 65.0%; Score 31; DB 4; Length 79;  
Best Local Similarity 62.5%; Pred. No. 40;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALMEQQHY 8  
Db 57 AVMEDRHY 64  
|:|:|:

RESULT 10  
US-09-543-681A-8198  
; Sequence 8198, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETTON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 8198  
; LENGTH: 164  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-8198

Query Match 66.0%; Score 31; DB 4; Length 164;  
Best Local Similarity 75.0%; Pred. No. 88;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LMEQQHYV 9  
Db 41 LNEMQHYV 48  
|:|:|:

RESULT 11  
US-09-489-039A-13178

```
; Sequence 13178, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13178
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13178

Query Match      66.0%; Score 31; DB 4; Length 235;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALMEQQHY 8
Db      94 AMLDRQHY 101

RESULT 12
US-09-248-796A-19806
; Sequence 19806, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19806
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19806

Query Match      66.0%; Score 31; DB 4; Length 393;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      4 EQQHYV 9
Db      207 QQQHYV 212

RESULT 13
US-09-489-039A-7205
; Sequence 7205, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7205
```

```
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7205

Query Match      66.0%; Score 31; DB 4; Length 478;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 ALMEQQHYV 9
Db      243 ALPENRHYV 251

RESULT 14
US-09-252-991A-24079
; Sequence 24079, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24079
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24079

Query Match      66.0%; Score 31; DB 4; Length 538;
Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 LMEQQHY 8
Db      100 LVEQQHF 106

RESULT 15
US-09-463-702A-2
; Sequence 2, Application US/09463702A
; Patent No. 6335435
; GENERAL INFORMATION:
; APPLICANT: AGENE Research Institute, Co., Ltd.
; APPLICANT: HIRAKI AND ASSOCIATES
; APPLICANT: SHIMAMOTO, AKIRO
; APPLICANT: KITAO, SAORI
; APPLICANT: FURUICHI, YASUHIRO
; TITLE OF INVENTION: HUMAN GENE RECO4 ENCODING HELICASE
; FILE REFERENCE: HIRAI150
; CURRENT APPLICATION NUMBER: US/09/463,702A
; CURRENT FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/JP98/03114
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: JAPAN 9/200387
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-463-702A-2

Query Match      65.0%; Score 31; DB 3; Length 1208;
Best Local Similarity 71.4%; Pred. No. 7.3e+02;
```



Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MEQQHVV 9  
|:|:|  
Db 358 MKQKHVV 364

Search completed: November 18, 2005, 09:35:57  
Job time : 15.6452 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 01:08:37 ; Search time 44.3226 Seconds  
(without alignments)  
84.961 Million cell updates/sec

Title: US-10-006-177-5

Perfect score: 47

Sequence: 1 ALMEQQHYV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	9	14 US-10-006-177-5	Sequence 5, Appli
2	47	100.0	769	10 US-09-984-130-67	Sequence 67, Appl
3	47	100.0	769	10 US-09-836-353A-67	Sequence 67, Appl
4	47	100.0	769	14 US-10-097-340-157	Sequence 157, App
5	47	100.0	769	15 US-10-295-027-480	Sequence 480, App
6	47	100.0	769	15 US-10-295-027-825	Sequence 825, App
7	47	100.0	769	15 US-10-295-027-845	Sequence 845, App
8	47	100.0	769	15 US-10-173-999-56	Sequence 56, Appl
9	47	100.0	769	20 US-11-050-926-157	Sequence 157, App
10	36	76.6	52	15 US-10-424-599-275477	Sequence 275477,
11	36	76.6	154	16 US-10-425-115-321251	Sequence 321251,

12	35	74.5	70	16	US-10-425-115-323882	Sequence 323882,
13	35	74.5	207	16	US-10-437-963-189508	Sequence 189508,
14	34	72.3	153	9	US-09-764-864-872	Sequence 872, App
15	34	72.3	294	16	US-10-408-765A-773	Sequence 773, App
16	33	70.2	96	16	US-10-437-963-112983	Sequence 112983,
17	33	70.2	229	9	US-09-866-562-92	Sequence 92, Appl
18	33	70.2	268	15	US-10-412-698B-730	Sequence 730, App
19	33	70.2	302	20	US-11-097-143-879	Sequence 879, App
20	33	70.2	623	15	US-10-425-114-59015	Sequence 59015, A
21	33	70.2	743	16	US-10-425-115-319142	Sequence 319142,
22	33	70.2	754	15	US-10-369-493-9797	Sequence 9797, App
23	33	70.2	775	10	US-09-934-455-450	Sequence 450, App
24	33	70.2	881	15	US-10-369-493-22252	Sequence 22252, A
25	32	68.1	95	16	US-10-767-701-33865	Sequence 33865, A
26	32	68.1	163	15	US-10-424-599-231149	Sequence 231149,
27	32	68.1	169	14	US-10-156-761-10091	Sequence 10091, A
28	32	68.1	241	15	US-10-369-493-21114	Sequence 21114, A
29	32	68.1	282	15	US-10-425-114-70396	Sequence 70396, A
30	32	68.1	382	17	US-10-732-923-2476	Sequence 2476, Ap
31	32	68.1	518	16	US-10-425-115-231624	Sequence 231624,
32	32	68.1	526	15	US-10-425-114-54307	Sequence 54307, A
33	32	68.1	591	17	US-10-755-087-18	Sequence 18, Appl
34	32	68.1	591	18	US-10-832-946-18	Sequence 18, Appl
35	32	68.1	634	20	US-11-097-143-41664	Sequence 41664, A
36	32	68.1	825	14	US-10-356-761-14417	Sequence 14417, A
37	32	68.1	974	16	US-10-437-963-127611	Sequence 127611,
38	32	68.1	1400	18	US-10-840-060-133	Sequence 133, App
39	32	68.1	1400	20	US-11-097-143-5457	Sequence 5457, Ap
40	31	66.0	51	16	US-10-425-115-357360	Sequence 357360,
41	31	66.0	56	15	US-10-424-599-180541	Sequence 180541,
42	31	66.0	92	16	US-10-425-115-254904	Sequence 254904,
43	31	66.0	104	16	US-10-437-963-195779	Sequence 195779,
44	31	66.0	106	16	US-10-425-115-276659	Sequence 276659,
45	31	66.0	123	16	US-10-425-115-232540	Sequence 232540,

#### ALIGNMENTS

##### RESULT 1

US-10-006-177-5  
; Sequence 5, Application US/10006177  
; Publication No. US20030165513A1  
; GENERAL INFORMATION:  
; APPLICANT: Ramakrishna, Venky  
; APPLICANT: Philip, Ramila  
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treatm  
; FILE REFERENCE: 26747-35  
; CURRENT APPLICATION NUMBER: US/10/006,177  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Epitopic Peptide  
US-10-006-177-5

Query Match 100.0%; Score 47; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQQHYV 9

Db 1 ALMEQQHYV 9

## RESULT 2

US-09-984-130-67  
 ; Sequence 67, Application US/09984130  
 ; Publication No. US20030055231A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ni et al.  
 ; TITLE OF INVENTION: 12 Human Secreted Proteins  
 ; FILE REFERENCE: PF489P2  
 ; CURRENT APPLICATION NUMBER: US/09/984,130  
 ; CURRENT FILING DATE: 2001-10-29  
 ; PRIOR APPLICATION NUMBER: 60/243,792  
 ; PRIOR FILING DATE: 2000-10-30  
 ; PRIOR APPLICATION NUMBER: 09/835,353  
 ; PRIOR FILING DATE: 2001-04-18  
 ; PRIOR APPLICATION NUMBER: 60/198,407  
 ; PRIOR FILING DATE: 2000-04-19  
 ; PRIOR APPLICATION NUMBER: PCT/US99/25031  
 ; PRIOR FILING DATE: 1999-10-27  
 ; PRIOR APPLICATION NUMBER: 60/105,971  
 ; PRIOR FILING DATE: 1998-10-28  
 ; NUMBER OF SEQ ID NOS: 149  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 67  
 ; LENGTH: 769  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-984-130-67

Query Match 100.0%; Score 47; DB 10; Length 769;

Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQQHYV 9

Db 662 ALMEQQHYV 670  
 |||||

## RESULT 3

US-09-836-353A-67  
 ; Sequence 67, Application US/09836353A  
 ; Publication No. US20030129685A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ni et al.  
 ; TITLE OF INVENTION: 12 Human Secreted Proteins  
 ; FILE REFERENCE: PF489P1  
 ; CURRENT APPLICATION NUMBER: US/09/836,353A  
 ; CURRENT FILING DATE: 2001-04-18  
 ; PRIOR APPLICATION NUMBER: 60/198,407  
 ; PRIOR FILING DATE: 2000-04-19  
 ; PRIOR APPLICATION NUMBER: PCT/US99/25031  
 ; PRIOR FILING DATE: 1999-10-27  
 ; PRIOR APPLICATION NUMBER: 60/105,971  
 ; PRIOR FILING DATE: 1998-10-28  
 ; NUMBER OF SEQ ID NOS: 147  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 67  
 ; LENGTH: 769  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-836-353A-67

Query Match 100.0%; Score 47; DB 10; Length 769;

Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQQHYV 9

Db 662 ALMEQQHYV 670  
 |||||

## RESULT 4

US-10-097-340-157  
 ; Sequence 157, Application US/10097340  
 ; Publication No. US20030087250A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: John MONAHAN  
 ; APPLICANT: Manjula GANNAVAPURU  
 ; APPLICANT: Sebastian HOERSCH  
 ; APPLICANT: Shubhangi KAMATKAR  
 ; APPLICANT: Steve G. KOVATS  
 ; APPLICANT: Rachel E. MEYERS  
 ; APPLICANT: Michael MORRISSEY  
 ; APPLICANT: Peter OLANDT  
 ; APPLICANT: Ami SEN  
 ; APPLICANT: Peter VEIBY  
 ; APPLICANT: Gordon B. MILLS  
 ; APPLICANT: Robert C. BAST, Jr.  
 ; APPLICANT: Karen LU  
 ; APPLICANT: Rosemarie SCHMANDT  
 ; APPLICANT: Xumei ZHAO  
 ; APPLICANT: Karen GLATT  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
 ; FILE REFERENCE: MRI-030  
 ; CURRENT APPLICATION NUMBER: US/10/097,340  
 ; CURRENT FILING DATE: 2002-03-14  
 ; PRIOR APPLICATION NUMBER: 60/276,025  
 ; PRIOR FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: 60/325,149  
 ; PRIOR FILING DATE: 2001-09-26  
 ; PRIOR APPLICATION NUMBER: 60/276,026  
 ; PRIOR FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: 60/324,967  
 ; PRIOR FILING DATE: 2001/09/26  
 ; PRIOR APPLICATION NUMBER: 60/311,732  
 ; PRIOR FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: 60/325,102  
 ; PRIOR FILING DATE: 2001-09-26  
 ; PRIOR APPLICATION NUMBER: 60/323,580  
 ; PRIOR FILING DATE: 2001-09-19  
 ; NUMBER OF SEQ ID NOS: 363  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 157  
 ; LENGTH: 769  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-097-340-157

Query Match 100.0%; Score 47; DB 14; Length 769;

Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQQHYV 9

Db 662 ALMEQQHYV 670  
 |||||

## RESULT 5

US-10-295-027-480  
 ; Sequence 480, Application US/10295027  
 ; Publication No. US20030232350A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Afar, Daniel  
 ; APPLICANT: Aziz, Natasha  
 ; APPLICANT: Ginsberg, Wendy M.  
 ; APPLICANT: Gish, Kurt C.  
 ; APPLICANT: Glynn, Richard  
 ; APPLICANT: Hevezi, Peter A.  
 ; APPLICANT: Mack, David H.  
 ; APPLICANT: Murray, Richard  
 ; APPLICANT: Watson, Susan R.  
 ; APPLICANT: Eos Biotechnology, Inc.  
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
 ; METHODS OF SCREENING FOR MODULATORS OF CANCER

; FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; PRIOR FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 480  
; LENGTH: 769  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-295-027-480

Query Match 100.0%; Score 47; DB 15; Length 769;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALMEQOHYV 9  
Db 662 ALMEQOHYV 670

RESULT 6  
US-10-295-027-825  
; Sequence 825, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Hevezi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 845  
; LENGTH: 769  
; TYPE: PRT  
; ORGANISM: Homo sapiens

; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 825  
; LENGTH: 769  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-295-027-825

Query Match 100.0%; Score 47; DB 15; Length 769;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALMEQOHYV 9  
Db 662 ALMEQOHYV 670

RESULT 7  
US-10-295-027-845  
; Sequence 845, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Hevezi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 845  
; LENGTH: 769  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-295-027-845

Query Match 100.0%; Score 47; DB 15; Length 769;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQOHV 9  
Db 662 ALMEQOHV 670

RESULT 8

US-10-173-999-56  
; Sequence 56, Application US/10173999  
; Publication No. US2004000563A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David H.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions  
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian  
; TITLE OF INVENTION: Cancer  
; FILE REFERENCE: 018501-002420US  
; CURRENT APPLICATION NUMBER: US/10/173,999  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/299,234  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: US 60/315,287  
; PRIOR FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/372,246  
; PRIOR FILING DATE: 2001-04-12  
; NUMBER OF SEQ ID NOS: 163  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 769  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-173-999-56

Query Match 100.0%; Score 47; DB 15; Length 769;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQOHV 9  
Db 662 ALMEQOHV 670

RESULT 9

US-11-050-926-157  
; Sequence 157, Application US/11050926  
; Publication No. US20050214831A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNAVAPU  
; APPLICANT: Sebastian HOERSCH  
; APPLICANT: Shubhangi KAWATKAR  
; APPLICANT: Steve G. KOVATS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU  
; APPLICANT: Rosemarie SCHMANDT  
; APPLICANT: Xumei ZHAO  
; APPLICANT: Karen GLATT  
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer

FILE REFERENCE: MRI-030  
; CURRENT APPLICATION NUMBER: US/11/050,926  
; CURRENT FILING DATE: 2005-02-04  
; PRIOR APPLICATION NUMBER: US/10/097,340  
; PRIOR FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001/09/26  
; PRIOR APPLICATION NUMBER: 60/311,732  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/325,102  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/323,580  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 157  
; LENGTH: 769  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-050-926-157

Query Match 100.0%; Score 47; DB 20; Length 769;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALMEQOHV 9  
Db 662 ALMEQOHV 670

RESULT 10

US-10-424-599-275477  
; Sequence 275477, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 275477  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_90777C.1.pap  
US-10-424-599-275477

Query Match 76.6%; Score 36; DB 15; Length 52;  
Best Local Similarity 75.0%; Pred. No. 20;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMEQOHV 9  
Db 14 LVEQOHL 21

RESULT 11

US-10-425-115-321251  
; Sequence 321251, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 321251  
; LENGTH: 154  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(154)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_56046C.1.pep  
US-10-425-115-321251

Query Match 76.6%; Score 36; DB 16; Length 154;  
Best Local Similarity 77.8%; Pred. No. 60;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALMEQOHYV 9  
Db 122 ALMQOQHCV 130

RESULT 12  
US-10-425-115-323882  
; Sequence 323882, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 323882  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_5844C.1.pep  
US-10-425-115-323882

Query Match 74.5%; Score 35; DB 16; Length 70;  
Best Local Similarity 62.5%; Pred. No. 42;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMEQOHYV 9  
Db 40 IMRQHYI 47

RESULT 13  
US-10-963-189508  
; Sequence 189508, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 189508  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_86010C.1.pep  
US-10-437-963-189508

Query Match 74.5%; Score 35; DB 16; Length 207;  
Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMEQOHYV 8  
Db 31 LVEQOHYV 37

RESULT 14  
US-09-764-864-872  
; Sequence 872, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 872  
; LENGTH: 153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-864-872

Query Match 72.3%; Score 34; DB 9; Length 153;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMEQOHYV 9  
Db 113 VMAQOHYV 120

RESULT 15  
US-10-408-765A-773  
; Sequence 773, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Watnick, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 773
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-773
```

```
Query Match      72.3%; Score 34; DB 16; Length 294;
Best Local Similarity 75.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy      2 LMEQQHYV 9
      :| |||||
Db      198 VMAQQHYV 205
```

```
Search completed: November 18, 2005, 10:44:34
Job time : 46.3226 secs
```



Result No.	Query No.	Score	Match	Length	DB	ID	Description	
1	46	100.0	9	5	ABG32306	Abg32306	HLA-A2 as	
2	46	100.0	100	8	ADs17294	Adsi17294	Human RPA	
3	46	100.0	440	5	ABp41244	Abp41244	Human ova	
4	46	100.0	616	5	AAm52617	Aam52617	Human rep	
5	46	100.0	616	6	AAo30065	Aao30065	Human rep	
6	41	89.1	616	5	AAu52618	Aau52618	Human rep	
7	38	82.6	237	6	ABU01081	Abu01081	S. pneumo	
8	37	80.4	915	7	ABo73598	AbO73598	Pseudomon	
9	36	78.3	48	8	ADm42984	Adm42984	Human INS	
10	36	78.3	48	8	ADM43042	Adm43042	Human INS	
11	36	78.3	349	2	AAy55934	Aay55934	Human ZC4	
12	36	78.3	399	5	ABp69243	Abp69243	Human pol	
13	36	78.3	959	8	ADr10462	Adr10462	Human pro	
14	36	78.3	1227	2	AAy55865	Aay55865	Full leng	
15	36	78.3	1500	8	ADM42398	Adm42398	Human INS	
16	36	78.3	1581	5	AAE24145	Aae24145	Human kin	
17	36	78.3	1582	8	ADO85780	Ado85780	Rac/axin/	
18	36	78.3	1589	8	ADM43056	Adm43056	Human INS	
19	35	76.1	238	6	ABU17208	Abu17208	Protein e	
20	35	76.1	254	6	ADA36364	Ada36364	Acinetoba	
21	35	76.1	384	6	ABu21264	Abu21264	Protein e	
22	34	73.9	202	8	ADO71150	Ado71150	Human org	
23	34	73.9	258	5	ABp65259	Abp65259	Bifidobac	
24	34	73.9	301	3	AAb50391	AbA50391	Human unc	
25	34	73.9	301	8	ADL61324	Adl61324	Human bol	

CC mammalian cell that can express the immunogen, is useful for inducing a  
 CC CTL response in vitro that is specific for a tumour cell expressing HLA-1  
 CC or HLA-2. The immunogen or cell is useful for inducing CTL for treating a  
 CC subject with cancer (carcinoma, preferably colorectal carcinoma, ovarian  
 CC carcinoma, lung carcinoma and prostate carcinoma). The immunogen is also  
 CC useful for screening and diagnostic agents, for gene screening in  
 CC patients afflicted with cancer, for screening a sample for the presence  
 CC of CTUs that specifically recognise the corresponding epitopes, as a  
 CC diagnostic tool to evaluate the efficacy of the immunotherapeutic  
 CC treatments, to prepare class I MHC (major histocompatibility class)  
 CC tetramers which are utilised in conjunction with flow cytometry to  
 CC quantitate the frequency of peptide-specific CTL that are present in a  
 CC samples of lymphocytes from an individual, and for stimulating the  
 CC production of antibodies for use in passive immunotherapy, for use as  
 CC diagnostic reagents, and for use as reagents in other processes such as  
 CC affinity chromatography. The present sequence is an immunogenic epitope  
 CC of the invention derived from human replication protein A

XX Sequence 9 AA;

Query Match 100.0%; Score 46; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9

Db 1 YLMDTSGKV 9

RESULT 2

ADSL17294

ID ADSL17294 standard; peptide; 100 AA.

XX AC ADSL17294;

XX DT 02-DEC-2004 (first entry)

XX DE Human RPA70 ssDNA-binding domain.

XX Single-stranded DNA binding protein; SSB; nucleic acid amplification;  
 KW nucleic acid engineering; ligase chain reaction;  
 KW polymerase chain reaction; PCR; recombination mediated cloning;  
 KW PCR-mediated gene replacement; mutagenesis; replication protein A; RPA70;  
 KW human.

XX OS Homo sapiens.

XX PN US2004I80342-A1.

XX PD 16-SEP-2004.

XX PF 11-MAR-2003; 2003US-00396575.

XX PR 11-MAR-2003; 2003US-00386575.

XX PA (REGC ) UNIV CALIFORNIA.

XX PI Haseltine CA, Kowalczykowski SC;

XX DR WPI; 2004-652518/63.

XX New multimer of Sulfolobus solfataricus single-stranded DNA-binding  
 PT protein, for use in performing nucleic acid amplification and  
 PT engineering.

XX Example 2; SEQ ID NO 7; 25pp; English.

XX The invention relates to Sulfolobus solfataricus protein multimers  
 CC comprising monomers of single-stranded DNA (ssDNA) binding protein (SSB).  
 CC The multimer is used in performing nucleic acid amplification and  
 CC engineering. The nucleic acid amplification is ligase chain reaction,  
 CC transcription-based amplification system, or self-sustained sequence  
 CC replication system, preferably a polymerase chain reaction (PCR). The

CC nucleic acid engineering is PCR-based DNA sequencing, recombination  
 CC mediated cloning, PCR-mediated gene replacement, PCR-mediated  
 CC recombination, reverse transcription (RT)-PCR cDNA synthesis, or in vitro  
 CC sequence mutagenesis. The present sequence is human RPA70 ssDNA-binding  
 CC domain.

SQ Sequence 100 AA;

Query Match 100.0%; Score 46; DB 8; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 0.17;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9

Db 47 YLMDTSGKV 55

RESULT 3

ABP41244

ID ABP41244 standard; protein; 440 AA.

XX AC ABP41244;

XX DT 23-AUG-2002 (first entry)

XX DE Human ovarian antigen HSUA20, SEQ ID NO:2376.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; reproductive system disorder; infertility;  
 KW pregnancy disorder; anovulation; polycystic ovary syndrome; PCOS;  
 KW ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive; chromosome 17p13.3.

XX OS Homo sapiens.

XX PN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US018569.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR N-PSDB; ABQ54321.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.

XX Claim 11; SEQ ID NO 2376; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovarian cancer and breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,

CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. Note: The  
 CC sequence represents a human ovarian antigen of the invention. The present  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 440 AA;

Query Match 100.0%; Score 46; DB 5; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 0.84;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
 |||||  
 Db 180 YLMDTSGKV 188

RESULT 4  
 AAM52617  
 ID AAM52617 standard; protein; 616 AA.  
 XX  
 XX AC AAM52617;  
 XX  
 XX DT 08-FEB-2002 (first entry)  
 XX DE Human replication protein A p70 subunit (RPA70).  
 XX  
 XX KW Human; replication protein A; p70 subunit; RPA70; unwindase; RPA;  
 KW single-stranded DNA-binding protein; HSSB; DNA replication; repair;  
 KW recombination; transcription; cell regulation; DNA polymerase; p53;  
 KW endonuclease recruitment; DNA metabolism; loss of heterozygosity; LOH;  
 KW cancer; tumour; allele-specific targeting; antisense therapy;  
 KW gene therapy.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 351  
 FT /note= "This residue is replaced by Ala in a polymorphic  
 FT variant (AAM52618)"  
 XX  
 XX PN US6309882-B1.  
 XX  
 XX PD 30-OCT-2001.  
 XX  
 XX PF 10-SEP-1999; 99US-00393529.  
 XX  
 XX PR 10-SEP-1999; 99US-00393529.  
 XX  
 XX PA (ISIS-) ISIS PHARM INC.  
 XX (VARI-) VARIAGENTS INC.  
 XX  
 XX PI Monia BP, Basilion JP, Stanton VP;  
 XX  
 XX DR WPI; 2002-040240/05.  
 XX N-PSDB; ABA02108.  
 XX  
 XX PT New antisense oligonucleotides targeted to nucleic acid encoding the  
 PT human replication protein A p70 subunit for diagnosing and treating  
 PT disease associated with expression of the subunit, including cancer.

XX

PS Example 2; Col 37-40; 24pp; English.

CC This sequence represents human replication protein A p70 subunit (RPA70).  
 CC The RPA70 subunit possesses unwindase activity and is able to bind to  
 CC single-stranded DNA and to certain double-stranded sequences with high  
 CC affinity. RPA70 is one of the three subunits of RPA (also known as human  
 CC single-stranded DNA-binding protein or HSSB) which is a single-stranded  
 CC DNA-binding protein that also interacts and regulates the function of a  
 CC number of other cellular proteins involved in DNA replication, repair,  
 CC recombination, transcription and cell regulation. The RPA70 subunit has  
 CC been shown to interact with DNA polymerase alpha during initiation of  
 CC replication and elongation, with polymerase delta, and with the tumour  
 CC suppressor protein p53, and also plays a role in recruiting endonucleases  
 CC involved in DNA repair. Due to the multiple roles of RPA70 in DNA  
 CC metabolism, loss of heterozygosity (LOH) at the RPA70 locus is associated  
 CC with many cancers. This makes RPA70 an attractive target for cancer  
 CC therapeutics based on allele-specific targeting, e.g., antisense  
 CC therapy. The invention relates to antisense oligonucleotides (ABA02110-  
 CC ABA02125) targetted to the human RPA70 gene, which inhibit its  
 CC expression. The oligonucleotides of the invention are useful for  
 CC diagnosis, prevention and treatment of conditions associated with RPA70,  
 CC particularly cancers, and can also be used in research reagents and kits

SQ Sequence 616 AA;

Query Match 100.0%; Score 46; DB 5; Length 616;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
 |||||  
 Db 347 YLMDTSGKV 355

RESULT 5  
 AAO30065  
 ID AAO30065 standard; protein; 616 AA.

XX  
 XX AC AAO30065;  
 XX  
 XX DT 03-SEP-2003 (first entry)  
 XX DE Human replication protein A (RPA) #1.  
 XX  
 XX KW Ataxia telangiectasia and Rad-3-related protein; ATR; cancer; ATRIP;  
 KW ATR-interaction protein; replication stress; cell-cycle checkpoint;  
 KW gene therapy; DNA damage; checkpoint signalling; replication protein A;  
 KW RPA; human.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX PN WO2003044214-A2.  
 XX  
 XX PD 30-MAY-2003.  
 XX  
 XX PF 20-NOV-2002; 2002WO-US037133.  
 XX  
 XX PR 20-NOV-2001; 2001US-0331821P.  
 XX  
 XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX  
 XX PI Elledge SJ, Cortez DK, Lee Z;  
 XX  
 XX DR WPI; 2003-457617/43.  
 XX N-PSDB; AAL60683.  
 XX  
 XX PT A composition for checkpoint signaling or for diagnosing and treating  
 PT cancer comprises an ataxia telangiectasia and Rad-3-related (ATR) and ATR  
 PT -interaction gene product polypeptide or polynucleotide.  
 XX  
 XX PS Disclosure; Page 158-160; 172pp; English.  
 XX

CC The invention relates to methods and compositions for checkpoint  
 CC signalling and for diagnosis and treatment of cancer. The composition  
 CC comprises an ataxia telangiectasia and Rad-3-related (ATR) and ATR-  
 CC interaction gene product polypeptide (ATRIP) or polynucleotide. The  
 CC invention is useful in checkpoint signalling, particularly signalling for  
 CC the presence of DNA damage or replication stress and activating cell-  
 CC cycle checkpoints. It is also used in diagnosing and treating cancer or  
 CC in screening for agents useful for diagnosing or treating cancer. The  
 CC invention is also used in gene therapy. The present sequence is human  
 CC replication protein A (RPA). This sequence is used to illustrate the  
 CC method of the invention  
 CC  
 XX SQ Sequence 616 AA;  
 CC  
 CC Query Match 100.0%; Score 46; DB 6; Length 616;  
 CC Best Local Similarity 100.0%; Pred. No. 1.2;  
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC 1 YLMDTSGKV 9  
 CC |||||  
 CC 347 YLMDTSGKV 355  
 CC  
 CC  
 CC RESULT 6  
 CC AAM52618  
 CC ID AAM52618 standard; protein; 616 AA.  
 CC XX  
 CC AC AAM52618;  
 CC  
 CC DT 08-FEB-2002 (first entry)  
 CC XX  
 CC DE Human replication protein A p70 subunit (RPA70) T351A variant.  
 CC XX  
 CC KW Human; replication protein A; p70 subunit; RPA70; unwindase; RPA;  
 CC KW single-stranded DNA-binding protein; HSSB; DNA replication; repair;  
 CC KW recombination; transcription; cell regulation; DNA polymerase; p53;  
 CC KW endonuclease recruitment; DNA metabolism; loss of heterozygosity; LOH;  
 CC KW cancer; tumour; allele-specific targeting; antisense therapy;  
 CC KW gene therapy; single nucleotide polymorphism; SNP.  
 CC XX  
 CC OS Homo sapiens.  
 CC XX  
 CC XX Key Location/Qualifiers  
 CC FT Misc-difference 351  
 CC FT /note= "This residue is replaced by Thr in a polymorphic  
 CC FT variant (AAM52617)"  
 CC XX  
 CC PN US6309882-B1.  
 CC XX  
 CC PD 30-OCT-2001.  
 CC XX  
 CC PF 10-SEP-1999; 99US-00393529.  
 CC XX  
 CC PR 10-SEP-1999; 99US-00393529.  
 CC XX  
 CC XX (ISIS-) ISIS PHARM INC.  
 CC PA (VARI-) VARIAGENICS INC.  
 CC XX  
 CC PI Monia BP, Basilion JP, Stanton VP;  
 CC XX  
 CC DR WPI; 2002-040240/05.  
 CC DR N-PSDB; ABA02109.  
 CC XX  
 CC PT New antisense oligonucleotides targeted to nucleic acid encoding the  
 CC PT human replication protein A p70 subunit for diagnosing and treating  
 CC PT disease associated with expression of the subunit, including cancer.  
 CC XX  
 CC XX Example 2; Page; 24pp; English.  
 CC PS  
 CC XX This sequence represents a polymorphic variant of the human replication  
 CC CC protein A p70 subunit (RPA70) in which the amino acid residue at position  
 CC CC 351 is Ala (rather than Thr), caused by an A to G nucleotide change at  
 CC CC position 1120 of the cDNA. The RPA70 subunit possesses unwindase activity

CC and is able to bind to single-stranded DNA and to certain double-stranded  
 CC sequences with high affinity. RPA70 is one of the three subunits of RPA  
 CC (also known as human single-stranded DNA-binding protein or HSSB) which  
 CC is a single-stranded DNA-binding protein that also interacts and  
 CC regulates the function of a number of other cellular proteins involved in  
 CC DNA replication, repair, recombination, transcription and cell  
 CC regulation. The RPA70 subunit has been shown to interact with DNA  
 CC polymerase alpha during initiation of replication and elongation, with  
 CC polymerase delta, and with the tumour suppressor protein p53, and also  
 CC plays a role in recruiting endonucleases involved in DNA repair. Due to  
 CC the multiple roles of RPA70 in DNA metabolism, loss of heterozygosity  
 CC (LOH) at the RPA70 locus is associated with many cancers. This makes  
 CC RPA70 an attractive target for cancer therapeutics based on allele-  
 CC specific targeting, e.g., antisense therapy. The invention relates to  
 CC antisense oligonucleotides (ABA02110-ABA02125) targeted to the human  
 CC RPA70 gene, which inhibit its expression. The oligonucleotides of the  
 CC invention are useful for diagnosis, prevention and treatment of  
 CC conditions associated with RPA70, particularly cancers, and can also be  
 CC used in research reagents and kits. Note: The present sequence is not  
 CC shown in the specification, but is derived from the RPA70 protein  
 CC sequence shown in columns 37-40 (AAM52617)  
 CC  
 XX SQ Sequence 616 AA;  
 CC  
 CC Query Match 89.1%; Score 41; DB 5; Length 616;  
 CC Best Local Similarity 88.9%; Pred. No. 13;  
 CC Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC 1 YLMDTSGKV 9  
 CC |||||  
 CC 347 YLMDASGKV 355  
 CC  
 CC  
 CC RESULT 7  
 CC ABU01081  
 CC ID ABU01081 standard; protein; 237 AA.  
 CC XX  
 CC AC ABU01081;  
 CC  
 CC DT 23-OCT-2003 (revised)  
 CC DT 11-FEB-2003 (first entry)  
 CC XX  
 CC XX S. pneumoniae type 4 strain protein from coding region #653.  
 CC XX  
 CC KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;  
 CC KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;  
 CC KW gene therapy; vaccine.  
 CC XX  
 CC OS Streptococcus pneumoniae; type 4 strain.  
 CC XX  
 CC PN WO200277021-A2.  
 CC XX  
 CC PD 03-OCT-2002.  
 CC XX  
 CC PF 27-MAR-2002; 2002WO-IB002163.  
 CC XX  
 CC XX 27-MAR-2001; 2001GB-00007658.  
 CC PR  
 CC XX (CHIR-) CHIRON SPA.  
 CC PA (GENO-) INST GENOMIC RES.  
 CC XX  
 CC XX Masignani V, Tettelin H, Fraser C;  
 CC XX  
 CC DR WPI; 2003-040579/03.  
 CC DR N-PSDB; ABX06365.  
 CC XX  
 CC PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,  
 CC PT useful as medicaments for treating or preventing a disease or infection  
 CC PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or  
 CC PT ear infection.  
 CC XX  
 CC PS Claim 1; SEQ ID NO 1306; 56pp; English.  
 CC XX

CC The invention relates to a protein comprising or having at least 50% identity to any of the 2469 amino acid sequences, identified in the specification (available on a computer readable format), or its fragment, expressed from 2469 of 2489 identified DNA coding regions from the Streptococcus pneumoniae type 4 strain genomic sequence appearing as ABS56454. Also included are an antibody which binds one of the proteins, treating a patient by administering the protein, DNA or antibody (in a composition), a kit comprising first and second primers, which are the nucleic acid cited above or fragments between nucleotides 8-100 of a sequence not defined in the specification, for amplifying a target sequence contained within a Streptococcus nucleic acid sequence, where the first primer is substantially complementary to the target sequence and the second primer is substantially complementary to the complement of the target sequence, and where the parts of the primers having substantial complementarity define the termini of the target sequence to be amplified, assay comprising contacting a test compound with the protein, and determining whether the test compound binds to the protein and a Streptococcus pneumoniae bacterium, where one or more genes encoding the proteins has been rendered inactive. The proteins, nucleic acid molecules, antibody and compositions are useful as medicaments for treating or preventing a disease or infection due to streptococcus bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis media or ear infection. They are also useful in developing vaccines, diagnostics and antibiotics. The methods are useful for identifying immunodominant proteins. The present sequence is one of the 2469 proteins expressed by the identified coding regions from the genomic sequence. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 23-Oct-2003 to standardise OS field)

XX CC Sequence 237 AA;

Query Match 82.6%; Score 38; DB 6; Length 237;  
Best Local Similarity 77.8%; Pred. No. 20;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
|||:|  
Db 38 YLMDTSGKV 46

RESULT 8  
ABO73598  
ID ABO73598 standard; protein; 915 AA.

XX AC ABO73598;

XX DT 29-JUL-2004 (first entry)

XX DE Pseudomonas aeruginosa polypeptide #5773.

XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX OS Pseudomonas aeruginosa.

XX PN US6551795-B1.

XX PD 22-APR-2003.

XX PF 18-FEB-1999; 99US-00252991.

XX PR 18-FEB-1998; 98US-0074788P.

XX PR 27-JUL-1998; 98US-0094190P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX FI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX DR WPI; 2003-615309/58.

XX DR N-PSDB; ABD07169.

XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

PT useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

XX PS Disclosure; SEQ ID NO 22344; 455pp; English.

XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biochip technology. Sequences ABO67826-ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

XX CC Sequence 915 AA;

Query Match 80.4%; Score 37; DB 7; Length 915;

Best Local Similarity 66.7%; Pred. No. 1.4e+02;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9

Db 402 YLLDTNGRV 410  
|||:|

RESULT 9

ADMA42984  
ID ADMA42984 standard; protein; 48 AA.

XX AC ADMA42984;

XX DT 03-JUN-2004 (first entry)

XX DE Human INSP082 exon 22 amino acid sequence SEQ ID NO:86.

XX KW human; INSP082; Germinal Centre Kinase subfamily; GCK subfamily;

XX KW STE20 protein kinase family; protein kinase; enzyme;

XX KW NCK-interacting kinase; NIK; NIK-like embryo specific kinase; NESK;

XX KW cytosolic; immunosuppressive; antiallergic; dermatological;

XX KW antiinflammatory; antiarthritic; CNS; endocrine; respiratory; cardiant;

XX KW analgesic; osteopathic; antimicrobial; gene therapy;

XX KW cell proliferative disorder; neoplasm; melanoma; solid tumour;

XX KW myeloproliferative disorder; leukaemia; non-Hodgkin lymphoma; leukopenia;

XX KW thrombocytopenia; angiogenesis disorder; Kaposi's sarcoma;

XX KW autoimmune disorder; inflammatory disorder; allergy;

XX KW inflammatory bowel disease; arthritis; psoriasis;

XX KW respiratory tract inflammation; asthma; organ transplant rejection;

XX KW cardiovascular disorder; hypertension; edema; angina; atherosclerosis;

XX KW thrombosis; sepsis; shock; reperfusion injury; ischaemia;

XX KW neurological disorder; central nervous system disease;

XX KW Alzheimer's disease; brain injury; amyotrophic lateral sclerosis; pain;

XX KW developmental disorder; metabolic disorder; diabetes mellitus;

XX KW osteoporosis; obesity; AIDS; renal disease; infection; spina bifida.

XX OS Homo sapiens.

XX PN WO2004022756-A2.

XX PD 18-MAR-2004.

XX PF 05-SEP-2003; 2003WO-GB003862.

XX PR 06-SEP-2002; 2002GB-00020770.

XX PA (ARES-) ARES TRADING SA.



CC other solid tumours; myeloproliferative disorders, such as leukaemia, non  
 CC -Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder,  
 CC Kaposi's sarcoma; autoimmune/inflammatory disorders, including allergy,  
 CC inflammatory bowel disease, arthritis, psoriasis and respiratory tract  
 CC inflammation, asthma, and organ transplant rejection; cardiovascular  
 CC disorders, including hypertension, edema, angina, atherosclerosis,  
 CC thrombosis, sepsis, shock, reperfusion injury, and ischaemia;  
 CC neurological disorders including central nervous system disease,  
 CC Alzheimer's disease, brain injury, amyotrophic lateral sclerosis, and  
 CC pain; developmental disorders; metabolic disorders including diabetes  
 CC mellitus, osteoporosis, and obesity; AIDS and renal disease; infections  
 CC including viral, bacterial, fungal and parasitic infections and other  
 CC pathological conditions and in particular developmental disorders of late  
 CC embryogenesis and neural tube defects such as spina bifida. The present  
 CC sequence represents the human INS091 exon 22 amino acid sequence, which  
 CC is used in the exemplification of the present invention.

XX Sequence 48 AA;

Query Match 78.3%; Score 36; DB 8; Length 48;  
 Best Local Similarity 87.5%; Pred. No. 9.1;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YLMDTSKG 8  
 |||||  
 Db 12 YLMDRSKG 19

RESULT 11

AAV55934  
 ID AAV55934 standard; protein; 349 AA.

XX AC AAV55934;

XX DT 18-FEB-2000 (first entry)  
 XX DE Human ZC4 protein.

KW Antirheumatic; antiarthritic; antiinflammatory; antiallergic; osteopathic;  
 KW antipsoriatic; antiarteriosclerotic; antiasthmatic; immunosuppressive;  
 KW neuroprotective; cardiac; cerebroprotective; cytotostatic; antidiabetic;  
 KW vulnery; STE20; protein kinase; STLK3; STLK4; STLK5; STLK6; STLK7;  
 KW ZC1; ZC2; ZC3; ZC4; KHS2; SULU1; SULU3; GEK2; PAK4; PAK5; antagonist;  
 KW antibody; gene therapy; rheumatoid arthritis; atherosclerosis; asthma;  
 KW inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;  
 KW rhinitis; autoimmunity; organ transplantation; multiple sclerosis;  
 KW myocardial infarction; cardiovascular disease; stroke; renal failure;  
 KW oxidative stress-related neurodegenerative disorder; Parkinson's disease;  
 KW amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;  
 KW ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;  
 KW mesangial disorder; growth regulation; wound healing; T cell activation;  
 KW immunosuppressant.

XX OS Homo sapiens.

XX PN WO9953036-A2.

XX PD 21-OCT-1999.

XX PF 13-APR-1999; 99WO-US008150.

XX PR 14-APR-1998; 98US-0081784P.

XX PA (SUGE-) SUGEN INC.

XX PI Plowman G, Martinez R, Whyte D;

XX DR WPI; 1999-611301/52.

XX DR N-PSDB; AAZ40486.

XX Novel kinase-related polypeptides used for the diagnosis and treatment of  
 PT kinase-related diseases and disorders.

XX

PS Claim 12; Page 283-284; 387pp; English.

XX This sequence represents a novel STE20-related protein kinase. The  
 CC invention relates to nucleic acid molecule encoding a kinase polypeptide  
 CC selected from STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3,  
 CC ZC4, KHS2, SULU1, SULU3, GEK2, PAK4 and PAK5. The proteins are used to  
 CC identify agonists and antagonists, and to raise antibodies. The  
 CC polynucleotides are useful in gene therapy protocols. The polynucleotides,  
 CC polypeptides, antibodies, antagonists and agonists may be used to treat  
 CC diseases such as immune-related disorders and diseases (e.g. rheumatoid  
 CC arthritis, atherosclerosis, chronic inflammatory bowel disease (e.g.  
 CC Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis,  
 CC rhinitis, autoimmunity, and organ transplantation, chronic inflammatory  
 CC pelvic disease, multiple sclerosis, organ transplantation, myocardial  
 CC infarction, cardiovascular disease, stroke, renal failure, oxidative  
 CC stress-related neurodegenerative disorders (e.g. amyotrophic lateral  
 CC sclerosis, Parkinson's disease and Leigh syndrome), cancer,  
 CC cardiomyopathies, ischemic disorders, inflammatory disorders, diabetes  
 CC mellitus, fibrotic and mesangial disorders. The proteins may also be  
 CC useful for cell growth regulation (e.g. in wound healing), T cell  
 CC activation, mitosis control, and as immunosuppressants  
 XX SQ Sequence 349 AA;

Query Match 78.3%; Score 36; DB 2; Length 349;  
 Best Local Similarity 87.5%; Pred. No. 77;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YLMDTSKG 8  
 |||||  
 Db 46 YLMDRSKG 53

RESULT 12

ABP69243  
 ID ABP69243 standard; protein; 399 AA.

XX AC ABP69243;

XX DT 20-JAN-2003 (first entry)

XX DE Human polypeptide SEQ ID NO 1290.

KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
 KW arthritis; cytotostatic; immunomodulator; nootropic; neuroprotective;  
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
 KW antiarthritic.

XX OS Homo sapiens.

XX PN WO200270539-A2.

XX PD 12-SEP-2002.

XX PF 05-MAR-2002; 2002WO-US005095.

XX PR 05-MAR-2001; 2001US-00799451.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao OA, Ren F;  
 XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,  
 XX Wehrman T, Wang J, Wang D, Drmanac RT;

XX DR WPI; 2002-759812/82.

XX DR N-PSDB; ABZ11460.

XX New polynucleotides comprising sequences assembled from expressed  
 PT sequence tags (ESTs), useful for treating cell-proliferative,



PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
 PT or coagulation disorders.

PS Claim 9; SEQ ID NO 1290; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a  
 CC nucleotide sequence selected from any of 948 sequences (AB211119-  
 CC AB212666) or their mature protein coding portion, active domain coding  
 CC protein or complementary sequences. The polynucleotides are useful for  
 CC identifying expressed genes or for physical mapping of human genome. The  
 CC encoded polypeptides (ABP68902-ABP68949) are useful as molecular weight  
 CC markers, as a food supplement, for generating antibodies, in medical  
 CC imaging, screening and diagnostic assays and for treating cell-  
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
 CC platelet or coagulation disorders, wound, burn, incision, ulcers, liver  
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
 CC arthritis, etc. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 399 AA;

Query Match 78.3%; Score 36; DB 5; Length 399;  
 Best Local Similarity 87.5%; Pred. No. 89;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 YLMDTSGK 8

Db 96 YLMDRSGK 103  
 |||||

RESULT 13

ADRI0462  
 ID ADRI0462 standard; protein; 959 AA.

XX AC ADRI0462;

XX DT 04-NOV-2004 (first entry)

XX DE Human protein useful for treating neurological disease Seq 3968.

XX KW human; oligo-capping method; diagnostic marker; gene therapy;  
 KW osteoporosis; neurological disease; Alzheimer's disease;  
 KW Parkinson's disease; dementia; short memory; cancer;  
 KW sense or motor function; emotional reaction; fear response; panic;  
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
 KW tranquiliser.

XX OS Homo sapiens.

XX PN EP1447413-A2.

XX PD 18-AUG-2004.

PF 12-FEB-2004; 2004EP-00003145.

XX 14-FEB-2003; 2003JP-00102207.

PR 09-MAY-2003; 2003JP-00131452.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;  
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;

XX WPI; 2004-583265/57.

DR N-PSDB; ADRI0423.

XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,  
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 3968; 2686pp; English.

XX This invention relates to novel, isolated full length human cDNA  
 CC molecules and the encoded proteins thereof. Specifically, it refers to  
 CC cDNA clones obtained by an oligo-capping method, where none of these  
 CC clones are identical to any known human mRNAs. The present invention  
 CC describes an immunoassay to identify agonists and antagonists, as well as  
 CC antibodies, antisense molecules and siRNAs that can all be used to bind  
 CC to and modulate expression of the cDNA molecules. As such, these  
 CC molecules are useful for diagnostic markers or therapeutic targets for  
 CC the various diseases or morbid states. In particular, they are useful in  
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
 CC disease, Parkinson's disease, dementia, short memory and various cancers,  
 CC as well as for maintaining equilibrium of sense or motor function, and  
 CC for treating emotional reaction, fear response and panic. Accordingly,  
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,  
 CC cytosstatic and tranquiliser activities. This polypeptide is a protein  
 CC encoded by a full length human cDNA sequence of the invention. NOTE: This  
 CC sequence is not given in the sequence listing of the specification but  
 CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-  
 CC office.

XX SQ Sequence 959 AA;

Query Match 78.3%; Score 36; DB 8; Length 959;  
 Best Local Similarity 87.5%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 YLMDTSGK 8

Db 899 YLMDRSGK 906  
 |||||

RESULT 14

AAAY55965  
 ID AAAY55965 standard; protein; 1227 AA.

XX AC AAAY55965;

XX DT 18-FEB-2000 (first entry)

XX DE Full length human ZC4 protein.

XX KW Antirheumatic; antiarthritic; antiinflammatory; antiallergic; osteopathic;  
 KW antipruritic; antiarteriosclerotic; antiasthmatic; immunosuppressive;  
 KW neuroprotective; cardiant; cerebroprotective; cytostatic; antidiabetic;  
 KW vulnery; STE20; protein kinase; STLK2; STLK3; STLK4; STLK5; STLK6; STLK7;  
 ZC1; ZC2; ZC3; ZC4; KHS2; SULU1; SULU3; GEK2; PAK4; PAK5; antagonist;  
 KW antibody; gene therapy; rheumatoid arthritis; arteriosclerosis; asthma;  
 KW inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;  
 KW rhinitis; autoimmunity; organ transplantation; multiple sclerosis;  
 KW myocardial infarction; cardiovascular disease; stroke; renal failure;  
 KW oxidative stress-related neurodegenerative disorder; Parkinson's disease;  
 KW amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;  
 KW ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;  
 KW mesangial disorder; growth regulation; wound healing; T cell activation;  
 KW immunosuppressant.

XX OS Homo sapiens.

XX WO9953036-A2.

XX 21-OCT-1999.

PF 13-APR-1999; 99WO-US008150.

PR 14-APR-1998; 98US-0081784P.

XX (SUGE-) SUGEN INC.

XX Plowman G, Martinez R, Whyte D;

XX WPI; 1999-611301/52.

DR N-PSDB; AAZ40539.



XX Novel kinase-related polypeptides used for the diagnosis and treatment of  
 PT kinase-related diseases and disorders.

PS Claim 11; Page 369-373; 387pp; English.

XX This sequence represents a novel STE20-related protein kinase. The  
 CC invention relates to nucleic acid molecule encoding a kinase polypeptide  
 CC selected from STUK2, STUK3, STUK4, STUK5, STUK6, STUK7, ZC1, ZC2, ZC3,  
 CC ZC4, KHS2, SUJ01, SUJ03, GSK2, PAK4 and PAK5. The proteins are used to  
 CC identify agonists and antagonists, and to raise antibodies. The  
 CC polynucleotides are useful in gene therapy protocols. The polynucleotides,  
 CC polypeptides, antibodies, antagonists and agonists may be used to treat  
 CC diseases such as immune-related disorders and diseases (e.g. rheumatoid  
 CC arthritis, atherosclerosis, chronic inflammatory bowel disease (e.g.  
 CC Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis,  
 CC rhinitis, autoimmunity, and organ transplantation, chronic inflammatory  
 CC pelvic disease, multiple sclerosis, organ transplantation, myocardial  
 CC infarction, cardiovascular disease, stroke, renal failure, oxidative  
 CC stress-related neurodegenerative disorders (e.g. amyotrophic lateral  
 CC sclerosis, Parkinson's disease and Leigh syndrome), cancer,  
 CC cardiomyopathies, ischemic disorders, inflammatory disorders, diabetes  
 CC mellitus, fibrotic and mesangial disorders. The proteins may also be  
 CC useful for cell growth regulation (e.g. in wound healing), T cell  
 CC activation, mitosis control, and as immunosuppressants

XX Sequence 1227 AA;

Query Match 78.3%; Score 36; DB 2; Length 1227;  
 Best Local Similarity 87.5%; Pred. No. 3e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSKG 8

Db 924 YLMDRSGK 931

RESULT 15

ADM42998

ID ADM42998 standard; protein; 1500 AA.

XX ADM42998;

DT 03-JUN-2004 (first entry)

XX Human INSP082 amino acid sequence SEQ ID NO:100.

XX human; INSP082; Germinal Centre Kinase subfamily; GCK subfamily;  
 KW STE20 protein kinase family; protein kinase; enzyme;  
 KW NCK-interacting kinase; NIK; NIK-like embryo specific kinase; NESK;  
 KW cytosolic; immunosuppressive; anti-allergic; dermatological;  
 KW anti-inflammatory; antiarthritic; CNS; endocrine; respiratory; cardiant;  
 KW analgesic; osteopathic; antimicrobial; gene therapy;  
 KW cell proliferative disorder; neoplasm; melanoma; solid tumour;  
 KW myeloproliferative disorder; leukaemia; non-Hodgkin lymphoma;  
 KW thrombocytopenia; angiogenesis disorder; Kaposi's sarcoma;  
 KW autoimmune disorder; inflammatory disorder; allergy;  
 KW inflammation; arthritis; psoriasis;  
 KW respiratory tract inflammation; asthma; organ transplant rejection;  
 KW cardiovascular disorder; hypertension; edema; angina; atherosclerosis;  
 KW thrombosis; sepsis; shock; reperfusion injury; ischaemia;  
 KW neurological disorder; central nervous system disease;  
 KW Alzheimer's disease; brain injury; amyotrophic lateral sclerosis; pain;  
 KW developmental disorder; metabolic disorder; diabetes mellitus;  
 KW osteoporosis; obesity; AIDS; renal disease; infection; spina bifida.

OS Homo sapiens.

XX WO2004022756-A2.

XX 18-MAR-2004.

XX 05-SEP-2003; 2003WO-GB003862.

XX 06-SEP-2002; 2002GB-00020770.

XX (ARES-) ARES TRADING SA.

XX Davids AR, Phelps CB, Fagan RJ, Power C, Yorke M;

XX WPI; 2004-248473/23.

XX N-PSDB; ADM42997.

XX New polypeptide that is a member of the Germinal Centre Kinase subfamily  
 PT of STE20 family of protein kinase, preferably NIK-like kinase or NIK-like  
 PT embryo specific kinase, useful for treating cancer, autoimmune or  
 PT inflammatory diseases.

XX Claim 4; SEQ ID NO 100; 121pp; English.

XX The present invention describes human INSP081, INSP082 and INSP091  
 CC proteins (I), which are members of the Germinal Centre Kinase (GCK)  
 CC subfamily of the STE20 family of protein kinase, preferably as NCK-  
 CC interacting kinases (NIKs) and more preferably as NIK-like embryo  
 CC specific kinases (NESKs). Also described: (1) purified nucleic acid  
 CC molecules (II) encoding (I); (2) a vector comprising (II); (3) a host  
 CC cell transformed with the vector; (4) a ligand that specifically binds to  
 CC the GCK subfamily of the STE20 family of protein kinases; (5) a compound  
 CC that increases or decreases the level of expression or activity of (I);  
 CC (6) a method of diagnosing a disease in a patient; (7) a pharmaceutical  
 CC composition comprising (I), (II), vector, host cell, ligand, or compound;  
 CC (8) a vaccine composition comprising (II); (9) a method for monitoring  
 CC the therapeutic treatment of disease in a patient; (10) a method for  
 CC identifying a compound that is effective in treating and/or diagnosing a  
 CC disease; (11) a transgenic or knockout non-human animal that has been  
 CC transformed to express lower, higher or absent levels of (I); and (12) a  
 CC method for screening a compound effective to treat a disease by  
 CC contacting the transgenic animal with a candidate compound and  
 CC determining the effect of the compound on the disease of the animal. (I)  
 CC has cytostatic, immunosuppressive, anti-allergic, dermatological,  
 CC anti-inflammatory, antiarthritic, CNS, endocrine, respiratory, cardiant,  
 CC analgesic, osteopathic and antimicrobial activities, and can be used in  
 CC gene therapy. The polypeptides (I), nucleic acid molecules (II), vectors,  
 CC host cells, ligands and methods of the invention can be used for treating  
 CC or diagnosing a disease, such as cell proliferative disorders, including  
 CC neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and  
 CC other solid tumours; myeloproliferative disorders, such as leukaemia, non-  
 CC Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder,  
 CC Kaposi's sarcoma; autoimmune/inflammatory disorders, including allergy,  
 CC inflammatory bowel disease, arthritis, psoriasis and respiratory tract  
 CC inflammation, asthma, and organ transplant rejection; cardiovascular  
 CC disorders, including hypertension, edema, angina, atherosclerosis,  
 CC thrombosis, sepsis, shock, reperfusion injury, and ischaemia;  
 CC neurological disorders including central nervous system disease,  
 CC Alzheimer's disease, brain injury, amyotrophic lateral sclerosis, and  
 CC pain; developmental disorders; metabolic disorders including diabetes  
 CC mellitus, osteoporosis, and obesity; AIDS and renal disease; infections  
 CC including viral, bacterial, fungal and parasitic infections and other  
 CC pathological conditions and in particular developmental disorders of late  
 CC embryogenesis and neural tube defects such as spina bifida. The present  
 CC sequence represents human INSP082, which is used in the exemplification  
 CC of the present invention.

XX Sequence 1500 AA;

Query Match 78.3%; Score 36; DB 8; Length 1500;

Best Local Similarity 87.5%; Pred. No. 3.7e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSKG 8

Db 1197 YLMDRSGK 1204

Search completed: November 18, 2005, 01:01:10

Job time : 52.7742 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 17, 2005, 22:11:06 ; Search time 8.90323 Seconds  
(without alignments)  
97.263 Million cell updates/sec

Title: US-10-006-177-6

Perfect score: 46

Sequence: 1 YLMDTSGKV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.\*

2: Pir2.\*

3: Pir3.\*

4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	616	1 A40457	replication protei
2	39	84.8	747	1 HIBPC7	internal virion pr
3	38	82.6	237	2 A95081	hypothetical prote
4	38	82.6	237	2 C97948	hypothetical prote
5	37	80.4	115	2 S73288	photosystem II pro
6	37	80.4	609	1 A43458	replication protei
7	37	80.4	633	2 B83478	probable two-compo
8	37	80.4	2825	2 T14271	Dcc4 protein, stre
9	36	78.3	133	2 AD2227	transposase all337
10	36	78.3	499	2 B96959	histidine kinase (
11	36	78.3	862	2 A05028	rpcC protein homol
12	35	76.1	573	2 B70047	two-component sens
13	34	73.9	87	2 S26282	retrovirus-related
14	34	73.9	147	2 AE2432	hypothetical prote
15	34	73.9	149	2 G72250	ribosomal protein
16	34	73.9	244	2 G82418	transcription regu
17	34	73.9	419	2 C69363	conserved hypothet
18	34	73.9	1876	2 T13801	phosphoinositide 3
19	33	71.7	185	2 D95115	thioredoxin family
20	33	71.7	191	2 H97984	conserved hypothet
21	33	71.7	240	2 D97762	probable ABC trans
22	33	71.7	372	2 AE3184	alcohol dehydrogen
23	33	71.7	591	2 C84220	propionyl-CoA carb
24	33	71.7	908	2 C70188	DNA polymerase I (
25	33	71.7	1021	2 E64576	hypothetical prote
26	33	71.7	1075	2 T47603	beta Galactosidase
27	33	71.7	1361	2 A29959	DNA-directed RNA p
28	32	69.6	133	2 AF2488	transposase alr708
29	32	69.6	135	2 H84523	probable thionin (

30 32 69.6 205 2 H96578  
31 32 69.6 214 2 JQ2001  
32 32 69.6 254 2 AB0855  
33 32 69.6 263 2 T39304  
34 32 69.6 376 2 JC7759  
35 32 69.6 420 2 AH3015  
36 32 69.6 425 2 G98268  
37 32 69.6 472 2 AH0109  
38 32 69.6 474 2 E83587  
39 32 69.6 601 2 F83471  
40 32 69.6 1161 2 E86626  
41 32 69.6 1433 2 T30261  
42 32 69.6 1440 2 C84639  
43 31.5 68.5 290 2 D69518  
44 31 67.4 80 2 C82669  
45 31 67.4 147 2 E46036

## ALIGNMENTS

## RESULT 1

A40457

replication protein A1 - human

N:Alternate names: replication protein A 70K chain

C:Species: Homo sapiens (man)

C&gt;Date: 28-Feb-1992 #sequence\_revision 27-Oct-1995 #text\_change 09-Jul-2004

C:Accession: A40457; A44501

R:Er dile, L.F.; Heyer, W.D.; Kolodner, R.; Kelly, T.J.

J. Biol. Chem. 266, 12090-12098, 1991

A:Title: Characterization of a cDNA encoding the 70-kDa single-stranded DNA-binding subu

A:Reference number: A40457; MUID:91268092; PMID:2050703

A:Accession: A40457

A:Molecule type: mRNA

A:Residues: 1-216, 'A', 218-616 &lt;ERD&gt;

A:Cross-references: UNIPROT:P27694; GB:M63488

A:Note: parts of the sequence determined by protein sequencing

A:Note: This sequence has been corrected in reference A44501

R:Er dile, L.F.; Heyer, W.D.; Kolodner, R.; Kelly, T.J.

J. Biol. Chem. 268, 2268, 1993

A:Title: Characterization of a cDNA encoding the 70-kDa single-stranded DNA-binding subu

A:Reference number: A44501; MUID:93131993; PMID:8420996

A:Accession: A44501

A:Molecule type: mRNA

A:Residues: 217 &lt;ER2&gt;

A:Cross-references: GB:M63488

A:Note: sequence correction

C:Genetics:

A:Gene: GDB:RPAL

A:Cross-references: GDB:138362; OMIM:179835

A:Map position: 17p13.3-17p13.3

C:Complex: Replication protein A is a trimer of 70K (A1), 32K (A2), and 14K (A3) chains.

C:Function:

A:Description: probable eukaryotic equivalent of prokaryotic single-stranded DNA-binding

Probably also has a role in the elongation stage of DNA replication

A:Pathway: DNA replication initiation

A:Note: the single-stranded DNA-binding activity resides in the A1 protein

C:Superfamily: replication protein A1

C:Keywords: DNA replication initiation; single-stranded DNA binding; trimer; zinc finger

F;481-503/Region: zinc finger CCCC motif

Query Match 100.0%; Score 46; DB 1; Length 616;

Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9

Dy 347 YLMDTSGKV 355

## RESULT 2

HIBPC7

internal virion protein C - phase T7

C;Species: phage T7  
C;Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 09-Jul-2004  
C;Accession: A04351; S42331  
R;Dunn, J.J.; Thompson, K.  
submitted to the Nucleic Acid Sequence Database, September 1982  
A;Reference number: A94615  
A;Accession: A04351  
A;Molecule type: DNA  
A;Residues: 1-747 <DUN>  
A;Cross-references: UNIPROT:P03725  
R;Dunn, J.J.; Studier, F.W.  
J. Mol. Biol. 166, 477-535, 1983  
A;Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genes  
A;Reference number: S42283; MUID:83241725; PMID:6864790  
A;Accession: S42331  
A;Molecule type: DNA  
A;Residues: 1-747 <DUN>  
A;Cross-references: EMBL:V01146; NID:9431187; PIDN:CAA24433.1; PID:g15609  
A;Note: the authors did not translate the codon for residue 1  
C;Genetics:  
A;Gene: 15  
A;Map position: 70.92-76.53  
C;Superfamily: phage T7 internal virion protein C

Query Match 84.8%; Score 39; DB 1; Length 747;  
Best Local Similarity 77.8%; Pred. No. 7.7; Mismatches 2; Indels 0; Gaps 0;  
Matches 7; Conservative 2

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 671 YLMDTTGQV 679

RESULT 3  
A95081  
hypothetical protein SP0698 [imported] - Streptococcus pneumoniae (strain TIGR4)  
C;Species: Streptococcus pneumoniae  
C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: A95081  
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A;Authors: Lofthus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A;Reference number: A95000; MUID:21357209; PMID:11463916  
A;Accession: A95081  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-237 <KUR>  
A;Cross-references: UNIPROT:Q97RU0; GB:AE005672; PIDN:AAK74842.1; PID:g14972173; GSPDB:C  
A;Experimental source: strain TIGR4  
C;Genetics:  
A;Gene: SP0698

Query Match 82.6%; Score 38; DB 2; Length 237;  
Best Local Similarity 77.8%; Pred. No. 3.8; Mismatches 1; Indels 0; Gaps 0;  
Matches 7; Conservative 1

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 38 YLMDTQGV 46

RESULT 4  
C97948  
hypothetical protein spr0611 [imported] - Streptococcus pneumoniae (strain R6)  
C;Species: Streptococcus pneumoniae  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.; E  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001  
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A;Reference number: A97872; MUID:21429245; PMID:111544234  
A;Accession: C97948  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-237 <KUR>  
A;Cross-references: UNIPROT:Q8CZ07; GB:AE007317; PIDN:AAK99415.1; PID:g15458194; GSPDB:G  
C;Genetics:  
A;Gene: spr0611

Query Match 82.6%; Score 38; DB 2; Length 237;  
Best Local Similarity 77.8%; Pred. No. 3.8; Mismatches 1; Indels 0; Gaps 0;  
Matches 7; Conservative 1

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 38 YLMDTQGV 46

RESULT 5  
S73268  
photosystem II protein W - red alga (Porphyra purpurea) chloroplast  
C;Species: chloroplast Porphyra purpurea  
C;Date: 19-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004  
C;Accession: S73268  
R;Reith, M.; Munholland, J.  
Plant Mol. Biol. Rep. 13, 333-335, 1995  
A;Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.  
A;Reference number: S73108  
A;Accession: S73268  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-115 <REI>  
A;Cross-references: UNIPROT:P51347; EMBL:U38804; NID:gl276652; PIDN:AAO8233.1; PID:gl27  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
C;Genetics:  
A;Gene: psbW  
A;Superfamily: photosystem II protein W  
C;Keywords: chloroplast; membrane-associated complex; photosynthesis; photosystem II; th

Query Match 80.4%; Score 37; DB 2; Length 115;  
Best Local Similarity 66.7%; Pred. No. 2.9; Mismatches 2; Indels 0; Gaps 0;  
Matches 6; Conservative 2

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 55 YLMDTEGOI 63

RESULT 6  
A43458  
replication protein A1 - African clawed frog  
N;Alternate names: replication protein A 70K chain  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 10-Jun-1993 #sequence\_revision 27-Oct-1995 #text\_change 09-Jul-2004  
C;Accession: A43458; S24638  
R;Adachi, Y.; Laemmli, U.K.  
J. Cell Biol. 119, 1-15, 1992  
A;Title: Identification of nuclear pre-replication centers poised for DNA synthesis in X  
A;Reference number: A43458; MUID:92407024; PMID:1527163  
A;Accession: A43458  
A;Molecule type: mRNA; protein  
A;Residues: 1-609 <ADA>  
A;Cross-references: UNIPROT:Q01588; EMBL:X67240; NID:g65065; PIDN:CAA47665.1; PID:g65066  
A;Experimental source: oocyte  
A;Note: sequence extracted from NCBI backbone (NCBIN:113390, NCBI:P:113391)  
A;Note: parts of the sequence determined by protein sequencing  
C;Complex: Replication protein A is a trimer of 70K (A1), 32K (A2), and 14K (A3) chains.  
C;Function:  
A;Description: thought to be the eukaryotic equivalent of prokaryotic single-stranded DN

A synthesis  
A:Pathway: DNA replication initiation  
C:Superfamily: replication protein A1  
C:Keywords: DNA replication initiation; single-stranded DNA binding; trimer; zinc finger  
F:472-494/Region: zinc finger CCCC motif

Query Match 80.4%; Score 37; DB 1; Length 609;  
Best Local Similarity 77.8%; Pred. No. 16;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLMDTSGKV 9  
338 HLMDSGKV 346  
:|||||

Db 1 YLMDTSGKV 9  
338 HLMDSGKV 346  
:|||||

RESULT 7  
B83478  
probable two-component sensor PA1336 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C:Accession: B83478  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olean, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: AB2950; MUID:20437337; PMID:10984043  
A:Accession: B83478  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-633 <STO>  
A:Cross-references: UNIPROT:Q9T408; GB:AE004563; GB:AE004091; NID:99947274; PIDN:AAG0472  
C:Genetics:  
A:Gene: PA1336  
C:Superfamily: signal transduction histidine kinase regulating C4-dicarboxylate transpor

Query Match 80.4%; Score 37; DB 2; Length 633;  
Best Local Similarity 66.7%; Pred. No. 17;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLMDTSGKV 9  
120 YLLDTNGRV 128  
|||:|:|

Db 1 YLMDTSGKV 9  
120 YLLDTNGRV 128  
|||:|:|

RESULT 8  
T14271  
Doc4 protein, stress-induced - mouse  
N:Alternate names: odz protein homolog  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: T14271  
R:Wang, X.Z.; Kuroda, M.; Sok, J.; Batchvarova, N.; Kimmel, R.; Chung, P.; Zinszner, H.;  
EMBO J. 17, 3619-3630, 1998  
A:Title: Identification of novel stress-induced genes downstream of chop.  
A:Reference number: Z17951; MUID:98315054; PMID:9649432  
A:Accession: T14271  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-2825 <WAN>  
A:Cross-references: UNIPROT:O70465; EMBL:AF059485; NID:g3170614; PID:g3170615; PIDN:AAC3  
C:Genetics:  
A:Gene: Doc4

Query Match 80.4%; Score 37; DB 2; Length 2825;  
Best Local Similarity 87.5%; Pred. No. 79;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YLMDTSGK 8  
1643 YLFDTSCK 1650  
|||:|:|

Db 1 YLMDTSGK 8  
1643 YLFDTSCK 1650  
|||:|:|

RESULT 9  
AD2227  
transposase all3371 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: AD2227  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AD2227  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-133 <KUR>  
A:Cross-references: UNIPROT:Q8YRS2; GB:BA000019; PIDN:BA075070.1; PID:g17132466; GSPDB:G  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all3371  
C:Superfamily: conserved hypothetical protein 4 (insertion sequence ISH1.8)

Query Match 78.3%; Score 36; DB 2; Length 133;  
Best Local Similarity 66.7%; Pred. No. 5.3;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YLMDTSGKV 9  
110 YMFDTGKV 118  
|:|:|:|

Db 1 YLMDTSGKV 9  
110 YMFDTGKV 118  
|:|:|:|

RESULT 10  
E96969  
histidine kinase (HAMP, Hiska and HATPase domains) [imported] - Clostridium acetobutylic  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: E96969  
R:Noelling, J.; Bretton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: E96969  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-499 <KUR>  
A:Cross-references: UNIPROT:Q97LJ5; GB:AB001437; PIDN:AAK78544.1; PID:g15023432; GSPDB:G  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC0565

Query Match 78.3%; Score 36; DB 2; Length 499;  
Best Local Similarity 77.8%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YLMDTSGKV 9  
102 YLDTTEGKV 110  
|||:|:|

Db 1 YLMDTSGKV 9  
102 YLDTTEGKV 110  
|||:|:|

RESULT 11  
A05028  
rpoC protein homolog - common tobacco chloroplast  
C:Species: Chloroplast Nicotiana tabacum (common tobacco)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004  
C:Accession: A05028  
R:Sugiura, M.  
submitted to the EMBL Data Library, August 1986  
A:Reference number: A00149  
A:Accession: A05028  
A:Molecule type: DNA  
A:Residues: 1-862 <SUG>

A;Cross-references: UNIPROT:P38550  
A;Experimental source: cv. Bright Yellow 4  
R;Shinozaki, K.; Ohme, M.; Tanaka, M.; Wakasugi, T.; Hayashida, N.; Matsubayashi, T.; Zeng, H.; Kamogashira, T.; Yamada, K.; Kusuha, F.; Takaiwa, F.; Kato, A.; Tondoh, N.; Shimada, J. 5, 2043-2049, 1986  
A;Title: The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization, sites, features  
A;Contents: annotation; gene organization, sites, features  
C;Genetics:  
A;Genome: chloroplast  
A;Start codon: TTC  
C;Superfamily: chloroplast DNA-directed RNA polymerase beta'-2 chain  
C;Keywords: chloroplast

Query Match 78.3%; Score 36; DB 2; Length 862;  
Best Local Similarity 66.7%; Pred. No. 37;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YLMDTSGKV 9  
|||||:|:  
Db 500 YLMDENGKI 508

RESULT 12  
B70047  
two-component sensor histidine kinase homolog yvrG - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: B70047  
R;Kunst, F.; Ogasawara, N.; Moser, J.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choudhry, A.; Ehrlich, S.D.; Emerson, P.T.; Encian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Authors: Schleicher, S.; Schroter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, T.; Winters, P.; Wipat, A.; Yamamoto, K.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshiyama, K.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: B70047  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-573 <KUN>  
A;Cross-references: UNIPROT:O34989; GB:Z99120; GB:Z99121; GB:AL009126; NID:G2635827; PID 168  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: yvrG

Query Match 76.1%; Score 35; DB 2; Length 573;  
Best Local Similarity 55.6%; Pred. No. 39;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLMDTSGKV 9  
|||||:|:  
Db 176 YLLDSNGKI 184

RESULT 13  
S26282  
retrovirus-related reverse transcriptase homolog (clone Wm7) - Welwitschia mirabilis  
C;Species: Welwitschia mirabilis  
C;Date: 19-Mar-1998 #sequence\_revision 19-Mar-1998 #text\_change 29-Jan-1999  
C;Accession: S26282  
R;Flavell, A.J.; Dunbar, E.; Anderson, R.; Pearce, S.R.; Hartley, R.; Kumar, A. Nucleic Acids Res. 20, 3639-3644, 1992  
A;Title: Tvl-copia group retrotransposons are ubiquitous and heterogeneous in higher plants  
A;Reference number: S26282; MUID:92350663; PMID:1379359

A;Accession: S26282  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-87 <FLA>  
A;Cross-references: EMBL:X65442  
A;Experimental source: clone Wm7  
A;Note: the sequence of residues 1-8, 83-87 and the corresponding nucleic acid sequence  
C;Genetics:  
A;Mobile element: retrotransposon copia-like Tyl  
C;Superfamily: retrovirus-related polypeptide

Query Match 73.9%; Score 34; DB 2; Length 87;  
Best Local Similarity 66.7%; Pred. No. 8.9;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YLMDTSGKV 9  
|||||:|:  
Db 69 YLMRTTGGI 77

RESULT 14  
AE2432  
hypothetical protein all5013 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. strain PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: AE2432  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S. DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AE2432  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-147 <KUR>  
A;Cross-references: UNIPROT:Q8YMC1; GB:BA000019; PIDN:BAB76712.1; PID:gl7134151; GSPDB:G1  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: all5013

Query Match 73.9%; Score 34; DB 2; Length 147;  
Best Local Similarity 75.0%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMDTSGKV 9  
|||||:|:  
Db 23 LIDTSGKI 30

RESULT 15  
G72250  
ribosomal protein L13 - Thermotoga maritima (strain MSB8)  
C;Species: Thermotoga maritima  
C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C;Accession: G72250  
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.  
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing  
A;Reference number: A72200; MUID:99287316; PMID:10360571  
A;Accession: G72250  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-149 <ARN>  
A;Cross-references: UNIPROT:Q9X1G5; GB:AE001797; GB:AE000512; NID:G4982017; PIDN:AAD3652  
C;Genetics:  
A;Gene: TMI454  
C;Superfamily: ribosomal protein L13

Query Match 73.9%; Score 34; DB 2; Length 149;

Best Local Similarity 66.7%; Pred. No. 16;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
|:| |||  
Db 23 YVVDASGRV 31

Search completed: November 18, 2005, 01:09:35  
Job time : 12.0032 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (C) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:04:26 ; Search time 39.9677 Seconds  
(without alignments)  
115.311 Million cell updates/sec

Title: US-10-006-177-6

Perfect score: 46

Sequence: 1 YLMDTSGKV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_Q3:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	616	1 REAL HUMAN	P27694 homo sapien
2	40	87.0	623	2 Q8VEE4	Q8VEE4 mus musculus
3	39	84.8	747	1 V1VC BPT7	P03725 bacterioph
4	39	84.8	747	2 Q6WY19	Q6WY19 bacterioph
5	39	84.8	747	2 Q6WY71	Q6WY71 bacterioph
6	39	84.8	747	2 Q6WYC5	Q6WYC5 bacterioph
7	39	84.8	747	2 Q6WYH4	Q6WYH4 bacterioph
8	39	84.8	747	2 Q858K0	Q858K0 yersinia pe
9	39	84.8	747	2 Q9T101	Q9T101 bacterioph
10	39	84.8	748	2 Q8W5U0	Q8W5U0 bacterioph
11	38	82.6	237	2 Q97RU0	Q97RU0 streptococc
12	38	82.6	237	2 Q8CZ07	Q8CZ07 streptococc
13	37	80.4	115	1 P8BW PORPU	P51347 porphyra pu
14	37	80.4	607	2 Q6NRZ5	Q6NRZ5 xenopus lae
15	37	80.4	609	1 REAL XENLA	Q01588 xenopus lae
16	37	80.4	633	2 Q91408	Q91408 pseudomonas
17	37	80.4	936	2 Q6CQK6	Q6CQK6 kluyveromyc
18	37	80.4	1198	2 Q80TF5	Q80TF5 mus musculus
19	37	80.4	2771	2 Q9WTS7	Q9WTS7 mus musculus
20	37	80.4	2825	2 Q70465	Q70465 mus musculus
21	36	78.3	133	2 Q8YRS2	Q8YRS2 anabaena sp
22	36	78.3	191	2 Q8CK50	Q8CK50 streptomyce
23	36	78.3	489	2 Q8M8U7	Q8M8U7 valeriana h
24	36	78.3	499	2 Q97LJ5	Q97LJ5 clostridium
25	36	78.3	959	2 Q5ZMP2	Q5ZMP2 homo sapien
26	36	78.3	1386	1 R5C2 OENHO	Q9mtm3 oenothera h
27	36	78.3	1387	2 Q68S16	Q68S16 panax ginseng
28	36	78.3	1388	1 R5C2 TOBAC	P38550 nicotiana t
29	36	78.3	1389	1 R5C2 ATREB	Q888Y1 atropa bell
30	36	78.3	1455	2 Q9R0G8	Q9R0G8 mus musculus
31	36	78.3	1455	2 Q9R0S4	Q9R0S4 mus musculus

32 36 78.3 1582 2 Q72ZY5  
33 36 78.3 1813 2 Q8ID65  
34 35 76.1 122 2 Q7VD72  
35 35 76.1 191 2 Q82PK6  
36 35 76.1 250 2 Q86YL8  
37 35 76.1 274 2 Q7UIS0  
38 35 76.1 286 2 Q7REK2  
39 35 76.1 456 2 Q74JV2  
40 35 76.1 573 2 Q34989  
41 35 76.1 4283 2 Q8UWL7  
42 34 73.9 147 2 Q8YMC1  
43 34 73.9 149 1 RL13 THEMA  
44 34 73.9 160 2 Q7URCS  
45 34 73.9 165 2 Q6P8P5

Q72zy5 homo sapien  
Q8id65 plasmodium  
Q7vd72 prochloroco  
Q82pk6 streptomyce  
Q86yl8 sulfolobus  
Q7uis0 rhodospirell  
Q7rek2 plasmodium  
Q74jv2 lactobacill  
Q34989 bacillus su  
Q8uwl7 fugu rubrip  
Q8ymc1 anabaena sp  
Q8ylg5 thermotoga  
Q7urcs rhodospirell  
Q6p8p5 mus musculu

#### ALIGNMENTS

RESULT 1  
ID REAL HUMAN STANDARD; PRT: 616 AA.  
AC P27694;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DE Replication protein A 70 kDa DNA-binding subunit (RP-A) (RP-A)  
DE (Replication factor-A protein 1) (Single-stranded DNA-binding protein).  
GN Name=RPAL; Synonyms=REPAL, RPA70;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=91268092; PubMed=2050703;  
RA Erdile L.F., Heyer W.-D., Kolodner R., Kelly T.J.;  
RT "Characterization of a cDNA encoding the 70-kDa single-stranded DNA-binding subunit of human replication protein A and the role of the protein in DNA replication."  
RL J. Biol. Chem. 266:12090-12098(1991).  
RN [2]  
RP REVISION TO 217.  
RX MEDLINE=93131993; PubMed=8420996;  
RA Erdile L.F., Heyer W.-D., Kolodner R., Kelly T.J.;  
RT "Type I human complement C2 deficiency. A 28-base pair gene deletion causes skipping of exon 6 during RNA splicing."  
RL J. Biol. Chem. 268:2268-2268(1993).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANT ALA-351.  
RA Livingston R.J., Rieder M.J., Chung M.-W., Ritchie T.K., Olson A.N., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D., Schackwitz W.S., Sherwood J.K., Sherwood A.M., Leithausen B.J., Nickerson D.A.;  
RT "NIHES-SNPs, environmental genome project, NIHES ES15478, Department of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Uterus;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Straubinger R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,



DE Replication protein A1.  
GN Name=Rpal;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Mammary tumor;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Loquellano N.J., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fanev J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schmerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Mammary tumor;  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC019119; AAH19119.1; --  
DR HSSP; P27694; 1FGU  
DR MGD; MGI:1915525; Rpal.  
DR GO; GO:0000793; C:condensed chromosome; IDA.  
DR GO; GO:0005634; C:nucleus; IDA.  
DR GO; GO:0003682; F:chromatin binding; IDA.  
DR GO; GO:0007126; P:meiosis; IDA.  
DR Pfam; PF04057; Rep-A\_N; 1.  
DR Pfam; PF01336; RNA\_anti; 1.  
DR TIGRFAMs; TIGR00617; Rpal; 1.  
SQ SEQUENCE 623 AA; 69037 MW; 2317F69F1E51B657 CRC64;

Query Match 87.0%; Score 40; DB 2; Length 623;  
Best Local Similarity 88.9%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 356 YLMDMSGKV 364

RESULT 3  
VIVC\_BPT7  
ID VIVC\_BPT7 STANDARD; PRT; 747 AA.  
AC P03725;  
DT 21-JUL-1996 (Rel. 01, Created)  
DT 21-JUL-1996 (Rel. 01, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Internal virion protein C.  
GN Name=15;  
OS Bacteriophage T7.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
OC T7-like viruses.  
OX NCBI\_TaxID=10760;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83241725; PubMed=6864790;  
RA Dunn J.J., Studier F.W.;

RT "Complete nucleotide sequence of bacteriophage T7 DNA and the  
RT locations of T7 genetic elements.";  
RL J. Mol. Biol. 166:477-535(1983).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; V01146; CAA24433.1; --  
DR PIR; A04351; HIBPC7.  
SQ SEQUENCE 747 AA; 84341 MW; 959C572B7B42C2B8 CRC64;

Query Match 84.8%; Score 39; DB 1; Length 747;  
Best Local Similarity 77.8%; Pred. No. 43;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 671 YLMDTTGQV 679

RESULT 4  
Q6WY19  
ID Q6WY19 PRELIMINARY; PRT; 747 AA.  
AC Q6WY19;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Gene 15.  
OS Bacteriophage T7.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
OC T7-like viruses.  
OX NCBI\_TaxID=10760;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2290536; PubMed=14629033;  
RA Bull J.J., Badgett M.R., Rokyta D., Molineux I.J.;  
RT "Experimental evolution yields hundreds of mutations in a functional  
RT viral genome.";  
RL J. Mol. Evol. 57:241-248(2003).  
DR EMBL; AY264778; AAP34162.1; --  
SQ SEQUENCE 747 AA; 84768 MW; E01495DFEA84095C CRC64;

Query Match 84.8%; Score 39; DB 2; Length 747;  
Best Local Similarity 77.8%; Pred. No. 43;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 671 YLMDTTGQV 679

RESULT 5  
Q6WY71  
ID Q6WY71 PRELIMINARY; PRT; 747 AA.  
AC Q6WY71;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Gene 15.  
OS Bacteriophage T7.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
OC T7-like viruses.  
OX NCBI\_TaxID=10760;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2290536; PubMed=14629033;  
RA Bull J.J., Badgett M.R., Rokyta D., Molineux I.J.;  
RT "Experimental evolution yields hundreds of mutations in a functional

```

RT viral genome."
RL J. Mol. Evol. 57:241-248(2003).
DR EMBL; AY264777; AAP34110.1; -.
SQ SEQUENCE 747 AA; 84681 MW; ASD41DF192BDF9CD CRC64;

Query Match      84.8%; Score 39; DB 2; Length 747;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 671 YLMDTTGQV 679

RESULT 6
Q6WYCS Q6WYCS PRELIMINARY; PRT; 747 AA.
AC Q6WYCS;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Gene 15.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22990536; PubMed=14629033;
RA Bull J.J., Badgett M.R., Rokyta D., Molineux I.J.;
RT "Experimental evolution yields hundreds of mutations in a functional
RT viral genome."
RL J. Mol. Evol. 57:241-248(2003).
DR EMBL; AY264776; AAP34056.1; -.
SQ SEQUENCE 747 AA; 84660 MW; 7EBBBE2845AF4F2 CRC64;

Query Match      84.8%; Score 39; DB 2; Length 747;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 671 YLMDTTGQV 679

RESULT 7
Q6WYH4 Q6WYH4 PRELIMINARY; PRT; 747 AA.
AC Q6WYH4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Gene 15.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22990536; PubMed=14629033;
RA Bull J.J., Badgett M.R., Rokyta D., Molineux I.J.;
RT "Experimental evolution yields hundreds of mutations in a functional
RT viral genome."
RL J. Mol. Evol. 57:241-248(2003).
DR EMBL; AY264775; AAP34007.1; -.
SQ SEQUENCE 747 AA; 84602 MW; AEDCF896B2EB3180 CRC64;

Query Match      84.8%; Score 39; DB 2; Length 747;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 671 YLMDTTGQV 679

```

```

Db 671 YLMDTTGQV 679

RESULT 8
Q858KO Q858KO PRELIMINARY; PRT; 747 AA.
AC Q858KO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Internal virion protein.
OS Yersinia pestis phage phiA1122.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=227720;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22803263; PubMed=12923098;
RX DOI=10.1128/JB.185.17.5248-5262.2003;
RA Garcia E., Elliott J.M., Ramanculov E., Chain P.S., Chu M.C.,
RA Molineux I.J.;
RT "The genome sequence of Yersinia pestis bacteriophage phiA1122 reveals
RT an intimate history with the coliphage T3 and T7 genomes."
RL J. Bacteriol. 185:5248-5262(2003).
DR EMBL; AY247822; AAP20538.1; -.
SQ SEQUENCE 747 AA; 84507 MW; 09C43BD10C1E08D8 CRC64;

Query Match      84.8%; Score 39; DB 2; Length 747;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 671 YLMDTTGQV 679

RESULT 9
Q9T101 Q9T101 PRELIMINARY; PRT; 747 AA.
AC Q9T101;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Internal virion protein C.
OS Bacteriophage phiYeO3-12.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=110457;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21122554; PubMed=11222590;
RX DOI=10.1128/JB.183.6.1928-1937.2001;
RA Rajunen M.I., Kiljunen S.J., Soederholm M.E.L., Skurnik M.;
RT "Complete genomic sequence of the lytic bacteriophage phiYeO3-12 of
RT Yersinia enterocolitica serotype O:3."
RL J. Bacteriol. 183:1928-1937(2001).
DR EMBL; AJ251805; CAB63636.1; -.
SQ SEQUENCE 747 AA; 85275 MW; AD7513D31519D57F CRC64;

Query Match      84.8%; Score 39; DB 2; Length 747;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 670 YLMDTTGQV 678

RESULT 10
Q8W5U0 Q8W5U0 PRELIMINARY; PRT; 748 AA.
ID Q8W5U0;
AC Q8W5U0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)

```

```

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Internal virion protein C.
GN Name=Gene 15;
OS Bacteriophage T3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10759;
RN [1]_TaxID=10759;
RP SEQUENCE FROM N.A.
RX STRAIN=Luria;
RC MEDLINE=22075331; PubMed=12079351; DOI=10.1016/S0022-2836(02)00384-4;
RA Pajunen M.I., Elizondo M.R., Skurnik M., Kielczawa J., Molineux I.J.;
RT "Complete nucleotide sequence and likely recombinatorial origin of
RT bacteriophage T3.";
RL J. Mol. Biol. 319:1115-1132(2002).
DR EMBL; AJ318471; CAC96303.1; -.
SQ SEQUENCE 748 AA; 84802 MW; 317A02C719AC8C99 CRC64;

Query Match 84.8%; Score 39; DB 2; Length 748;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 672 YLMDTSGKV 680
|||||:|
|||||:|

RESULT 11
Q97RU0 PRELIMINARY; PRT; 237 AA.
AC Q97RU0;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SP0698.
GN OrderedLocustNames=SP0698;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E.K., Khouri H.M., Wolf A.M., Uitterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL; AE007378; AAK74842.1; -.
DR PIR; A95081; A95081.
DR TIGR; SP0698; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 237 AA; 26927 MW; E4DCB74A1825DD CRC64;

Query Match 82.6%; Score 38; DB 2; Length 237;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 38 YLMDTSGKV 46
|||||:|
|||||:|

RESULT 12
Q8CZ07 PRELIMINARY; PRT; 237 AA.
AC Q8CZ07;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein spr0611.
GN OrderedLocustNames=spr0611;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmore R., Glaes J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Niclas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008440; AAK99415.1; -.
DR PIR; C97948; C97948.
KW Complete proteome.
SQ SEQUENCE 237 AA; 26835 MW; B575D0660D4B9323 CRC64;

Query Match 82.6%; Score 38; DB 2; Length 237;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 38 YLMDTSGKV 46
|||||:|
|||||:|

RESULT 13
PSBW_PORPU STANDARD; PRT; 115 AA.
ID PSBW_PORPU
AC PS1347;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Photosystem II reaction center W protein.
GN Name=psbw;
OS Porphyra purpurea.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reith M.E., Munnholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC CC -1- FUNCTION: Subunit of the water oxidation complex of photosystem II
CC reaction center complex.
CC -1- SIMILARITY: Belongs to the psbw family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U38804; AAC08233.1; -.

```

DR DR DR

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9

Db :|||:||||  
338 HLMDSSGKV 346

Search completed: November 18, 2005, 01:08:10  
Job time : 45.9677 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 00:48:52 ; Search time 13.6452 Seconds  
(without alignments)  
49.237 Million cell updates/sec

Title: US-10-006-177-6

Perfect score: 46

Sequence: 1 YLMDTSGKV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pap:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pap:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pap:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pap:\*
- 5: /cgn2\_6/ptodata/1/iaa/PTCUS.COMB.pap:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	616	3	US-09-393-529-2
2	46	100.0	616	4	US-09-396-149-7
3	37	80.4	609	4	US-09-396-149-6
4	37	80.4	915	4	US-09-252-991A-22344
5	36	78.3	349	4	US-09-688-188B-16
6	36	78.3	349	4	US-09-291-417D-16
7	36	78.3	1227	4	US-09-688-188B-105
8	36	78.3	1227	4	US-09-291-417D-105
9	35	76.1	254	4	US-09-328-352-7651
10	34	73.9	301	4	US-09-949-016-6865
11	34	73.9	333	4	US-09-949-016-6865
12	34	73.9	1876	2	US-08-609-049A-12
13	34	73.9	1876	2	US-08-609-049A-12
14	34	73.9	1876	3	US-09-170-996-28
15	34	73.9	1876	3	US-09-170-996-28
16	33	71.7	185	4	US-09-583-110-5222
17	33	71.7	194	4	US-09-107-433-5104
18	33	71.7	529	3	US-09-433-248A-6
19	32	69.6	207	4	US-09-248-796A-17685
20	32	69.6	288	4	US-09-902-540-12765
21	32	69.6	694	4	US-09-252-991A-22481
22	31	67.4	82	1	US-08-370-225-19
23	31	67.4	82	1	US-08-461-859-19
24	31	67.4	82	5	PCT-US93-10069-19
25	31	67.4	92	4	US-08-943-667-30
26	31	67.4	97	4	US-09-583-110-3949
27	31	67.4	97	4	US-09-107-433-4804

## ALIGNMENTS

### RESULT 1

US-09-393-529-2  
; Sequence 2, Application US/09393529  
; Patent No. 6309882  
; GENERAL INFORMATION:  
; APPLICANT: Monia, Brett P  
; TITLE OF INVENTION: Antisense Modulation of Replication  
; FILE OF INVENTION: Protein A p70 Subunit  
; FILE REFERENCE: ISPH-0383  
; CURRENT APPLICATION NUMBER: US/09/393.529  
; CURRENT FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-393-529-2

Query Match 100.0%; Score 46; DB 3; Length 616;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YLMDTSGKV 9  
|||  
Db 347 YLMDTSGKV 355

### RESULT 2

US-09-396-149-7  
; Sequence 7, Application US/09396149  
; Patent No. 6538176  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; TITLE OF INVENTION: Maize Replication Protein A and Use  
; FILE REFERENCE: 5718-59  
; CURRENT APPLICATION NUMBER: US/09/396.149  
; CURRENT FILING DATE: 1999-09-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-396-149-7

Query Match 100.0%; Score 46; DB 4; Length 616;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

28 31 67.4 121 4 US-09-134-000C-4262 Sequence 4262, Ap  
29 31 67.4 153 4 US-09-287-070-2 Sequence 2, Appli  
30 31 67.4 162 4 US-09-538-092-55 Sequence 55, Appl  
31 31 67.4 168 4 US-09-252-991A-18629 Sequence 18629, A  
32 31 67.4 182 2 US-08-721-925A-3 Sequence 3, Appli  
33 31 67.4 335 4 US-09-902-540-15208 Sequence 15208, A  
34 31 67.4 452 4 US-09-463-712C-8 Sequence 8, Appli  
35 31 67.4 452 4 US-09-865-415-6 Sequence 6, Appli  
36 31 67.4 490 1 US-08-294-770A-2 Sequence 2, Appli  
37 31 67.4 491 2 US-08-448-735C-2 Sequence 2, Appli  
38 31 67.4 556 2 US-09-016-000-4 Sequence 4, Appli  
39 31 67.4 556 4 US-09-156-793D-2 Sequence 2, Appli  
40 31 67.4 556 4 US-08-943-667-1 Sequence 1, Appli  
41 31 67.4 638 4 US-09-252-991A-27646 Sequence 27646, A  
42 31 67.4 915 4 US-09-949-016-7528 Sequence 7528, Ap  
43 31 67.4 1245 2 US-08-801-263A-3 Sequence 3, Appli  
44 31 67.4 1245 2 US-08-801-263A-6 Sequence 6, Appli  
45 31 67.4 1245 2 US-08-801-263A-10 Sequence 10, Appli

```
Qy      1 YLMDTSGKV 9
      |||||
Db      347 YLMDTSGKV 355

RESULT 3
US-09-396-149-6
; Sequence 6, Application US/09396149
; Patent No. 6538176
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Replication Protein A and Use
; FILE REFERENCE: 5718-59
; CURRENT APPLICATION NUMBER: US/09/396.149
; CURRENT FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-396-149-6

      Query Match      80.4%; Score 37; DB 4; Length 609;
      Best Local Similarity 77.8%; Pred. No. 24;
      Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 YLMDTSGKV 9
      :|||:||||
Db      338 HLMDSGKV 346

RESULT 4
US-09-252-991A-22344
; Sequence 22344, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22344
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22344

      Query Match      80.4%; Score 37; DB 4; Length 915;
      Best Local Similarity 66.7%; Pred. No. 37;
      Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 YLMDTSGKV 9
      |||:|:|
Db      402 YLLDTNGRV 410

RESULT 5
US-09-688-188B-16
; Sequence 16, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
```

```
; CURRENT APPLICATION NUMBER: US/09/688.188B
; CURRENT FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-188B-16

      Query Match      78.3%; Score 36; DB 4; Length 349;
      Best Local Similarity 87.5%; Pred. No. 21;
      Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 YLMDTSGK 8
      |||||
Db      46 YLMDRSCK 53

RESULT 6
US-09-291-417D-16
; Sequence 16, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291.417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-291-417D-16

      Query Match      78.3%; Score 36; DB 4; Length 349;
      Best Local Similarity 87.5%; Pred. No. 21;
      Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 YLMDTSGK 8
      |||||
Db      46 YLMDRSCK 53

RESULT 7
US-09-688-188B-105
; Sequence 105, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/09/688.188B
; CURRENT FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
```

```

; LENGTH: 1227
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-188B-105

Query Match      78.3%; Score 36; DB 4; Length 1227;
Best Local Similarity 87.5%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGK 8
Db 924 YLMDRSGK 931

RESULT 8
US-09-291-417D-105
; Sequence 105, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 1227
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-291-417D-105

Query Match      78.3%; Score 36; DB 4; Length 1227;
Best Local Similarity 87.5%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGK 8
Db 924 YLMDRSGK 931

RESULT 9
US-09-328-352-7651
; Sequence 7651, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7651
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7651

Query Match      76.1%; Score 35; DB 4; Length 254;
Best Local Similarity 87.5%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMDTSGKV 9
Db 197 LVDTSKV 204

RESULT 10

```

```

US-09-949-016-6865
; Sequence 6865, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6865
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6865

Query Match      73.9%; Score 34; DB 4; Length 301;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 138 YEMETSGKI 146

RESULT 11
US-09-949-016-9952
; Sequence 9952, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9952
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9952

Query Match      73.9%; Score 34; DB 4; Length 333;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 170 YEMETSGKI 178

RESULT 12
US-08-609-049A-12
; Sequence 12, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.

```

APPLICANT: Molz, Lisa  
APPLICANT: Chen, Yen-Wen  
TITLE OF INVENTION: No. 5948664e1 PI 3-Kinase Polypeptides  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,049A  
FILING DATE: 29-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 2307K-0637000US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1876 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-609-049A-12

Query Match 73.9%; Score 34; DB 2; Length 1876;  
Best Local Similarity 77.8%; Pred. No. 3.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
|:|||||  
Db 557 YMDTSIKV 565

RESULT 13  
US-08-609-049A-28  
Sequence 28, Application US/08609049A  
Patent No. 5948664  
GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Molz, Lisa  
APPLICANT: Chen, Yen-Wen  
TITLE OF INVENTION: No. 5948664e1 PI 3-Kinase Polypeptides  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,049A  
FILING DATE: 29-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.

REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 2307K-0637000US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1876 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-609-049A-28

Query Match 73.9%; Score 34; DB 2; Length 1876;  
Best Local Similarity 77.8%; Pred. No. 3.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
|:|||||  
Db 557 YMDTSIKV 565

RESULT 14  
US-09-170-996-12  
Sequence 12, Application US/09170996  
Patent No. 6291220  
GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Molz, Lisa  
APPLICANT: Chen, Yen-Wen  
TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/170,996  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609,049  
FILING DATE: 29-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 2307K-0637000US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1876 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-170-996-12

Query Match 73.9%; Score 34; DB 3; Length 1876;  
Best Local Similarity 77.8%; Pred. No. 3.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
|:|||||

Db 557 YMDTSIKV 565

RESULT 15

US-09-170-996-28  
 ; Sequence 28, Application US/09170996  
 ; Patent No. 6291220  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Williams, Lewis T.  
 ; APPLICANT: Molz, Lisa  
 ; APPLICANT: Chen, Yen-Wen  
 ; TITLE OF INVENTION: No. 6291220el PI 3-Kinase Polypeptides  
 ; NUMBER OF SEQUENCES: 32  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, 8th Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/170,996  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/609,049  
 ; FILING DATE: 29-FEB-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Dow, Karen B.  
 ; REGISTRATION NUMBER: 29,684  
 ; REFERENCE/DOCKET NUMBER: 2307K-063700US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-326-2400  
 ; TELEFAX: 415-326-2422  
 ; INFORMATION FOR SEQ ID NO: 28:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1876 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-170-996-28

Query Match 73.9%; Score 34; DB 3; Length 1876;  
 Best Local Similarity 77.8%; Pred. No. 3.3e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YLMDTSIKV 9

|:|||||

Db 557 YMDTSIKV 565

Search completed: November 18, 2005, 09:35:58  
 Job time : 14.6452 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 01:08:37 ; Search time 44.3226 Seconds  
(without alignments)  
84.961 Million cell updates/sec

Title: US-10-006-177-6

Perfect score: 46

Sequence: 1 YLMDTSGKV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	46	100.0	9 14	US-10-006-177-6
2	46	100.0	100 16	US-10-386-575-7
3	46	100.0	229 20	US-11-035-703-13
4	46	100.0	440 15	US-10-264-049-2376
5	46	100.0	616 14	US-10-372-686-7
6	46	100.0	616 14	US-10-371-558-7
7	46	100.0	616 14	US-10-300-453A-41
8	46	100.0	616 14	US-10-375-553-7
9	46	100.0	616 15	US-10-372-553-7
10	38	82.6	237 17	US-10-472-928-1306
11	37	80.4	229 20	US-11-035-703-14

Sequence 6, Appli  
Sequence 7, Appli  
Sequence 13, Appli  
Sequence 2376, Ap  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 41, Appli  
Sequence 7, Appli  
Sequence 1306, Ap  
Sequence 14, Appli

12	37	80.4	609	14	US-10-372-686-6	Sequence 6, Appli
13	37	80.4	609	14	US-10-371-558-6	Sequence 6, Appli
14	37	80.4	609	14	US-10-375-553-6	Sequence 6, Appli
15	37	80.4	609	15	US-10-372-553-6	Sequence 6, Appli
16	37	80.4	2771	9	US-09-808-602-82	Sequence 82, Appli
17	37	80.4	2771	10	US-09-800-198-70	Sequence 70, Appli
18	36	78.3	349	10	US-09-291-417-16	Sequence 16, Appli
19	36	78.3	349	16	US-10-725-329-16	Sequence 16, Appli
20	36	78.3	349	18	US-10-725-121-16	Sequence 16, Appli
21	36	78.3	1227	10	US-09-291-417-105	Sequence 105, App
22	36	78.3	1227	16	US-10-725-329-105	Sequence 105, App
23	36	78.3	1227	18	US-10-725-121-105	Sequence 105, App
24	36	78.3	1539	18	US-10-840-512-207	Sequence 207, App
25	36	78.3	1581	15	US-10-415-011-16	Sequence 16, Appli
26	35	76.1	101	16	US-10-437-963-144707	Sequence 144707,
27	35	76.1	191	14	US-10-156-761-8434	Sequence 8434, Ap
28	35	76.1	238	15	US-10-282-122A-45132	Sequence 45132, Ap
29	35	76.1	384	15	US-10-282-122A-49188	Sequence 49188, A
30	34	73.9	132	15	US-10-424-599-198393	Sequence 198393,
31	34	73.9	225	17	US-10-876-542-7	Sequence 7, Appli
32	34	73.9	229	16	US-10-437-963-105543	Sequence 105543,
33	34	73.9	301	14	US-10-116-255-33	Sequence 33, Appli
34	34	73.9	301	16	US-10-648-593-248	Sequence 248, App
35	34	73.9	301	17	US-10-732-923-19176	Sequence 19176, A
36	34	73.9	301	17	US-10-732-923-19190	Sequence 19190, A
37	34	73.9	301	17	US-10-732-923-19351	Sequence 19351, A
38	34	73.9	301	17	US-10-732-923-19352	Sequence 19352, A
39	34	73.9	301	18	US-10-756-149-4948	Sequence 4948, App
40	34	73.9	419	15	US-10-369-493-998	Sequence 998, App
41	34	73.9	454	20	US-11-097-143-6111	Sequence 6111, Ap
42	34	73.9	907	16	US-10-437-963-154788	Sequence 154788,
43	34	73.9	1876	20	US-11-097-143-12702	Sequence 12702, A
44	34	73.9	1876	20	US-11-097-143-27420	Sequence 27420, A
45	34	73.9	2758	18	US-10-467-535-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1  
US-10-006-177-6  
; Sequence 6, Application US/10006177  
; Publication No. US20030165513A1  
; GENERAL INFORMATION:  
; APPLICANT: Ramakrishna, Venky  
; APPLICANT: Ross, Mark  
; APPLICANT: Philip, Ramila  
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treatm  
; TITLE OF INVENTION: Diagnosis of Cancer  
; FILE REFERENCE: 26747-35  
; CURRENT APPLICATION NUMBER: US/10/006,177  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR FILING DATE: US/60/251,022  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: US/60/256,824  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Epitopic Peptide  
US-10-006-177-6

Query Match 100.0%; Score 46; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9

Db 1 YLMDTSGKV 9

```
RESULT 2
US-10-386-575-7
; Sequence 7, Application US/10386575
; Publication No. US20040180342A1
; GENERAL INFORMATION:
; APPLICANT: Haseltine, Cynthia A.
; APPLICANT: Kowalczykowski, Stephen C.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Multimers of S. solfataricus Single-Stranded
; TITLE OF INVENTION: DNA-Binding Protein and Methods of Use Thereof
; FILE REFERENCE: 023070-125300US
; CURRENT APPLICATION NUMBER: US/10/386.575
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (438)
; OTHER INFORMATION: replication protein A 70 kDa DNA-binding subunit
; OTHER INFORMATION: (RPA70) ssDNA-binding domain
US-10-386-575-7

Query Match          100.0%; Score 46; DB 16; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
   |||||
Db 47 YLMDTSGKV 55

RESULT 3
US-11-035-703-13
; Sequence 13, Application US/11035703
; Publication No. US20050164265A1
; GENERAL INFORMATION:
; APPLICANT: Kowalczykowski, Stephen C.
; APPLICANT: Chedin, Frederic
; APPLICANT: Seitz, Erica M.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Single Stranded DNA Binding Proteins From Archaea and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: 023070-097910US
; CURRENT APPLICATION NUMBER: US/11/035.703
; CURRENT FILING DATE: 2005-01-14
; PRIOR APPLICATION NUMBER: US/09/631.616
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/147,680
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (438)
; OTHER INFORMATION: RPA70 198-427
US-11-035-703-13

Query Match          100.0%; Score 46; DB 20; Length 229;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
   |||||
Db 149 YLMDTSGKV 157

RESULT 4
US-10-372-686-7
; Sequence 7, Application US/10372686
; Publication No. US20030159185A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D
; CURRENT APPLICATION NUMBER: US/10/372.686
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-372-686-7

Query Match          100.0%; Score 46; DB 14; Length 616;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
   |||||
Db 347 YLMDTSGKV 355

RESULT 6
US-10-371-558-7
; Sequence 7, Application US/10371558
; Publication No. US20030163840A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
```

```
US-10-264-049-2376
; Sequence 2376, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264.049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2376
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (438)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-2376

Query Match          100.0%; Score 46; DB 15; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
   |||||
Db 180 YLMDTSGKV 188

RESULT 5
US-10-372-686-7
; Sequence 7, Application US/10372686
; Publication No. US20030159185A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D
; CURRENT APPLICATION NUMBER: US/10/372.686
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-372-686-7

Query Match          100.0%; Score 46; DB 14; Length 616;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
   |||||
Db 347 YLMDTSGKV 355

RESULT 6
US-10-371-558-7
; Sequence 7, Application US/10371558
; Publication No. US20030163840A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
```



; TITLE OF INVENTION: Replication Protein A and Use  
; FILE REFERENCE: 0899D2  
; CURRENT APPLICATION NUMBER: US/10/371,558  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/396,149  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/123,896  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 60/100,690  
; PRIOR FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-371-558-7

Query Match 100.0%; Score 46; DB 14; Length 616;  
Best Local Similarity 100.0%; Pred. No. 1.7; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 347 YLMDTSGKV 355

RESULT 7  
US-10-300-453A-41  
; Sequence 41, Application US/10300453A  
; Publication No. US20030165934A1  
; GENERAL INFORMATION:  
; APPLICANT: ELLEDGE, STEPHEN J.  
; APPLICANT: CORTEZ, DAVID K.  
; APPLICANT: ZOU, LEE  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN CHECKPOINT SIGNALING  
; FILE REFERENCE: P02339US  
; CURRENT APPLICATION NUMBER: US/10/300,453A  
; CURRENT FILING DATE: 2003-03-12  
; PRIOR APPLICATION NUMBER: 60/331,821  
; PRIOR FILING DATE: 2001-11-20  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 41  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-300-453A-41

Query Match 100.0%; Score 46; DB 14; Length 616;  
Best Local Similarity 100.0%; Pred. No. 1.7; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 347 YLMDTSGKV 355

RESULT 8  
US-10-375-553-7  
; Sequence 7, Application US/10375553  
; Publication No. US20030174240A1  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; TITLE OF INVENTION: Replication Protein A and Use  
; FILE REFERENCE: 0899D3  
; CURRENT APPLICATION NUMBER: US/10/375,553  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/396,149  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/123,896  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 60/100,690

; PRIOR FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-375-553-7

Query Match 100.0%; Score 46; DB 14; Length 616;  
Best Local Similarity 100.0%; Pred. No. 1.7; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 347 YLMDTSGKV 355

RESULT 9  
US-10-372-553-7  
; Sequence 7, Application US/10372553  
; Publication No. US20040098769A1  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; TITLE OF INVENTION: Replication Protein A and Use  
; FILE REFERENCE: 0899D3  
; CURRENT APPLICATION NUMBER: US/10/372,553  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/396,149  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/123,896  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 60/100,690  
; PRIOR FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-372-553-7

Query Match 100.0%; Score 46; DB 15; Length 616;  
Best Local Similarity 100.0%; Pred. No. 1.7; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 1 YLMDTSGKV 9  
| | | | |  
Db 347 YLMDTSGKV 355

RESULT 10  
US-10-472-928-1306  
; Sequence 1306, Application US/10472928  
; Publication No. US20050020813A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SpA  
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026926WO  
; CURRENT APPLICATION NUMBER: US/10/472,928  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: GB-0107658.7  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 4979  
; SOFTWARE: Seqwin99, version 1.03  
; SEQ ID NO 1306  
; LENGTH: 237  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; OTHER INFORMATION: hypothetical protein  
; OTHER INFORMATION: Cellular location: membrane

```
; OTHER INFORMATION: Similar to strain R6 sequence 15902655 (e-127)
US-10-472-928-1306

Query Match      82.6%; Score 38; DB 17; Length 237;
Best Local Similarity 77.8%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 38 YLMEOTGKV 46
      |||:||||

RESULT 11
US-11-035-703-14
; Sequence 14, Application US/11035703
; Publication No. US20050164265A1
; GENERAL INFORMATION:
; APPLICANT: Kowalczykowski, Stephen C.
; APPLICANT: Chedin, Frederic
; APPLICANT: Seltz, Erica W.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Single Stranded DNA Binding Proteins From Archaea and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: 023070-097910US
; CURRENT APPLICATION NUMBER: US/11/035,703
; CURRENT FILING DATE: 2005-01-14
; PRIOR APPLICATION NUMBER: US/09/631,616
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/147,680
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Xenopus laevis
; FEATURE:
; OTHER INFORMATION: RPA70 190-418
US-11-035-703-14

Query Match      80.4%; Score 37; DB 20; Length 229;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 149 HLMDSSGKV 157
      |||:||||

RESULT 12
US-10-372-686-6
; Sequence 6, Application US/10372686
; Publication No. US20030159185A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D
; CURRENT APPLICATION NUMBER: US/10/372,686
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-372-686-6

Query Match      80.4%; Score 37; DB 14; Length 609;
Best Local Similarity 77.8%; Pred. No. 11e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 338 HLMDSSGKV 346
      |||:||||

RESULT 13
US-10-371-558-6
; Sequence 6, Application US/10371558
; Publication No. US20030163840A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D2
; CURRENT APPLICATION NUMBER: US/10/371,558
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-371-558-6

Query Match      80.4%; Score 37; DB 14; Length 609;
Best Local Similarity 77.8%; Pred. No. 11e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 338 HLMDSSGKV 346
      |||:||||

RESULT 14
US-10-375-553-6
; Sequence 6, Application US/10375553
; Publication No. US20030174240A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D3
; CURRENT APPLICATION NUMBER: US/10/375,553
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-375-553-6

Query Match      80.4%; Score 37; DB 14; Length 609;
Best Local Similarity 77.8%; Pred. No. 11e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9
Db 338 HLMDSSGKV 346
      |||:||||
```

RESULT 15  
US-10-372-553-6  
; Sequence 6, Application US/10372553  
; Publication No. US20040098769A1  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; TITLE OF INVENTION: Replication Protein A and Use  
; FILE REFERENCE: 0899D3  
; CURRENT APPLICATION NUMBER: US/10/372,553  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/396,149  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/123,896  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 60/100,690  
; PRIOR FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Xenopus laevis  
US-10-372-553-6

Query Match 80.4%; Score 37; DB 15; Length 609;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YLMDTSGKV 9  
:|||||  
Db 338 HLMDSGKV 346

Search completed: November 18, 2005, 10:44:37  
Job time : 47.3226 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: .November 17, 2005, 22:03:06 ; Search time 48.7742 Seconds  
(without alignments)  
71.366 Million cell updates/sec

Title: US-10-006-177-7

Perfect score: 48

Sequence: 1 ILDDIGHGV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	9	ABG32307	HLA-A2 as
2	48	100.0	397	AD78151	Human ext
3	48	100.0	401	AAW15225	Abl inter
4	48	100.0	446	AAW15225	Abl inter
5	48	100.0	475	AAW15225	Abl inter
6	48	100.0	486	AAW15225	Abl inter
7	48	100.0	486	AAW15225	Abl inter
8	48	100.0	543	AAW15225	Abl inter
9	46	95.8	173	AAW15225	Abl inter
10	46	95.8	232	AAW15225	Abl inter
11	46	95.8	241	AAW15225	Abl inter
12	46	95.8	367	AAW15225	Abl inter
13	46	95.8	394	AAW15225	Abl inter
14	46	95.8	496	AAW15225	Abl inter
15	43	89.6	365	AAW15225	Abl inter
16	43	89.6	366	AAW15225	Abl inter
17	43	89.6	370	AAW15225	Abl inter
18	43	89.6	421	AAW15225	Abl inter
19	43	89.6	436	AAW15225	Abl inter
20	43	89.6	476	AAW15225	Abl inter
21	43	89.6	476	AAW15225	Abl inter
22	40	83.3	473	AAW15225	Abl inter
23	40	83.3	552	AAW15225	Abl inter
24	39	81.2	380	AAW15225	Abl inter
25	39	81.2	629	AAW15225	Abl inter

26	38	79.2	356	6	ABM67210	Phototrab
27	38	79.2	362	7	ABO71333	Pseudomon
28	38	79.2	483	7	ABO74278	Pseudomon
29	37	77.1	286	4	AAU47798	Propionib
30	37	77.1	286	6	ABM44317	Propionib
31	36	75.0	147	4	AAU53509	Propionib
32	36	75.0	147	6	ABM50028	Propionib
33	36	75.0	255	7	ABO62558	Klebsiell
34	36	75.0	299	7	ADF05697	Bacterial
35	36	75.0	416	5	ABR38842	A. niger
36	36	75.0	465	7	ABO80641	Pseudomon
37	36	75.0	677	7	ABO71660	Pseudomon
38	35	72.9	68	8	ABO54965	Human gen
39	35	72.9	114	4	AAU07666	Rainbow t
40	35	72.9	217	7	ABO73959	Pseudomon
41	35	72.9	321	7	ABO75487	Pseudomon
42	35	72.9	417	4	ABG96724	Putative
43	35	72.9	420	4	ABB71794	Drosophil
44	35	72.9	422	8	ADN48206	Thermococ
45	35	72.9	433	5	ABP28150	Streptoco

## ALIGNMENTS

### RESULT 1

ABG32307

ID ABG32307 standard; peptide; 9 AA.

XX AC ABG32307;

XX DT 05-NOV-2002 (first entry)

XX DE HLA-A2 associated immunogenic peptide from human Abl binding protein 3.

XX KW Human; immunogen; epitope; HLA-A1; human leukocyte antigen; CTL;

XX KW cytotoxic Y lymphocyte; cytostatic; cancer; colorectal carcinoma;

XX KW ovarian carcinoma; lung carcinoma; prostate carcinoma; vaccine; tumour;

XX KW HLA-2; passive immunotherapy; Abl binding protein 3.

XX OS Homo sapiens.

XX PN WO200246416-A2.

XX PD 13-JUN-2002.

XX PF 04-DEC-2001; 2001WO-US047290.

XX PR 04-DEC-2000; 2000US-0251022P.

XX PA 20-DEC-2000; 2000US-0256824P.

XX PI (ARGO-) ARGONEX INC.

XX PI Ramakrishna V, Ross M, Philip R;

XX DR WPI; 2002-619021/66.

XX PT New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte,

XX PT and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.

XX FS Claim 1; Page 50; 60pp; English.

XX CC The invention relates to an immunogen comprising an isolated polypeptide

XX CC whose amino acid sequence comprises an epitopic peptide, does not include

XX CC MAGE 4 or MFG-E8 proteins, or consists of MAGE D protein or its

XX CC immunologically active fragment. Also included are a polynucleotide

XX CC encoding the immunogen or its complement, a vector comprising the

XX CC polynucleotide, a mammalian cell comprising the vector and expressing the

XX CC polynucleotide, a vaccine composition comprising the immunogen and an

XX CC antibody specific for the immunogen. The immunogen is useful for inducing

XX CC a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour

XX CC cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is

XX CC useful for inducing a CTL response when administered to a subject. A



PN WO9715587-A1.  
PD 01-MAY-1997.  
XX 25-OCT-1996; 96WO-US017047.  
XX 27-OCT-1995; 95US-00549004.  
XX (UYDU-) UNIV DUKE.  
XX Pendergast A, Dai Z;  
XX WPI; 1997-258949/23.  
DR N-PSDB; AAT66904.  
XX Human Abl interactor protein - useful for identifying potential cancer  
PT chemotherapeutic compounds.  
PS Claim 3; Fig 1A; 92pp; English.  
XX This polypeptide sequence comprises the human Abl interactor (Abl)  
CC protein Abi-2 (AAW15225), which interacts with both the SH3 domain and C-  
CC terminal sequences of c-Abl protein tyrosine kinase, and which has a dual  
CC role as a regulator and effector of c-Abl, and functions as a tumour  
CC suppressor in cells. The amino acid sequence was deduced from a nucleic  
CC acid (AAT66904) isolated from a human lymphocyte library using a yeast  
CC two-hybrid system with the Abl SH3 domain as bait. Abi-2 and its  
CC fragments can be produced in host cells using a claimed method. It can be  
CC used to generate Abi-2 specific antibodies and to screen for compounds  
CC that modulate Abl oncogenic activity for potential use as cancer  
CC chemotherapeutic agents  
XX Sequence 401 AA;  
SQ Query Match 100.0%; Score 48; DB 2; Length 401;  
Best Local Similarity 100.0%; Pred. No. 1.9; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ILDDIGHGV 9  
Db 100 ILDDIGHGV 108  
RESULT 4  
AAB64369  
ID AAB64369 standard; protein; 446 AA.  
XX AAB64369;  
XX 22-MAR-2001 (first entry)  
XX Amino acid sequence of human intracellular signalling molecule INTRAL.  
XX Human; intracellular signalling molecule; INTRA; immunosuppressive;  
KW cytosolic; neuroprotective; nootropic; antiarteriosclerotic; cancer;  
KW antiinflammatory; anti-HIV; neuroleptic; antibacterial; antifungal;  
KW antiviral; antiparasitic; antihelminthic; antiparkinsonian; AIDS;  
KW cell proliferative disorder; arteriosclerosis autoimmune; epilepsy;  
KW inflammatory disorder; Addison's disease; gastrointestinal disorder;  
KW neurological disorder; Parkinson's disease; Creutzfeldt-Jakob disease;  
KW mental disorder; schizophrenia; anxiety.  
XX Homo sapiens.  
OS WO200077040-A2.  
XX WO200077040-A2.  
XX 21-DEC-2000.  
XX 16-JUN-2000; 2000WO-US016636.  
XX 16-JUN-1999; 99US-0139566P.  
XX 17-AUG-1999; 99US-0149640P.  
XX 09-NOV-1999; 99US-0164417P.

XX (INCY-) INCYTE GENOMICS INC.  
XX Yue H, Tang YT, Hillman JL, Lal P, Bandman O, Baughn MR;  
PI Azimzai Y, Yang J, Reddy R, Lu DAM;  
XX WPI; 2001-025334/03.  
DR N-PSDB; AAF32638.  
XX New human intracellular signaling molecules, useful for the diagnosis,  
PT prevention and treatment of cell proliferative, autoimmune, inflammatory,  
PT neurological, gastrointestinal, reproductive and developmental disorders.  
XX Claim 5; Page 112-113; 192pp; English.  
XX Sequences AAF32638 - AAF32689 represent cDNA encoding human intracellular  
CC signalling molecules INTRAL - INTRA52, represented in AAB64369 -  
CC AAB64420. Modulators of the intracellular signalling molecules of the  
CC invention exhibit immunosuppressive; cytostatic; neuroprotective;  
CC nootropic; antiarteriosclerotic; antiinflammatory; anti-HIV; neuroleptic;  
CC antibacterial; antifungal; antiviral; antiparasitic; antihelminthic; and  
CC antiparkinsonian activity. INTRA polypeptides their agonists and  
CC antagonists are useful for the treatment of a condition associated with  
CC decreased or increased expression of functional INTRA. Disorders  
CC associated with abnormal INTRA expression or activity include cell  
CC proliferative disorders e.g. arteriosclerosis and cancer; autoimmune or  
CC inflammatory disorders e.g. Addison's disease and acquired  
CC immunodeficiency syndrome (AIDS); viral, bacterial, fungal, parasitic,  
CC protozoal and helminthic infections; gastrointestinal disorders e.g.  
CC dysphagia and irritable bowel syndrome; neurological disorders e.g.  
CC epilepsy and Parkinson's disease; prion diseases e.g. Creutzfeldt-Jakob  
CC disease and mental disorders e.g. anxiety, schizophrenia and Tourette's  
CC disorder. Antibodies immuno specific for the INTRA proteins may also be  
CC useful in the diagnosis of the above disorders  
XX Sequence 446 AA;  
SQ Query Match 100.0%; Score 48; DB 4; Length 446;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ILDDIGHGV 9  
Db 145 ILDDIGHGV 153  
RESULT 5  
ADE63221  
ID ADE63221 standard; protein; 475 AA.  
XX ADE63221;  
XX 29-JAN-2004 (first entry)  
XX Human Protein NP\_005750, SEQ ID NO 9158.  
XX Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX Homo sapiens.  
OS WO2003016475-A2.  
XX WO2003016475-A2.  
XX 27-FEB-2003.  
XX 14-AUG-2002; 2002WO-US025765.  
XX 14-AUG-2001; 2001US-0312147P.  
XX 01-NOV-2001; 2001US-0346382P.  
XX 26-NOV-2001; 2001US-0333347P.  
XX (GEHO ) GEN HOSPITAL CORP.

PA (PAB ) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 DR GENBANK; NP\_005750.  
 XX  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 XX Claim 1; Page; 1017pp; English.  
 XX  
 CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 475 AA;  
 Query Match 100.0%; Score 48; DB 7; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 2.2; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 0;  
 QY 1 ILDDIGHGV 9  
 DB 145 ILDDIGHGV 153  
 RESULT 6  
 ADR66705  
 ID ADR66705 standard; protein; 486 AA.  
 XX  
 AC ADR66705;  
 XX  
 XX 02-DEC-2004 (first entry)  
 DT  
 XX Human prostatic carcinoma derived protein SEQ ID 217 #3.  
 DE  
 XX human; cytostatic; diagnosis; prostatic cancer;  
 KW differential expression analysis.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2004076614-A2.  
 PN  
 XX 10-SEP-2004.  
 PD  
 XX 22-FEB-2004; 2004WO-DE000433.  
 PF  
 XX 27-FEB-2003; 2003DE-01009985.  
 PR

PR 14-MAY-2003; 2003DE-01022134.  
 XX  
 PA (HINZ/) HINZMANN B.  
 PA (DAHL/) DAHL E.  
 PA (ROSE/) ROSENTHAL A.  
 PA (HERM/) HERMANN K.  
 PA (PILA/) PILARSKY C.  
 XX  
 XX Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;  
 PI Schmitt A, Beckmann G, Bruemendorf T, Kimmern H, Roepcke S;  
 PI Xinzhang L, Staub E;  
 XX WPI; 2004-653386/63.  
 DR  
 XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,  
 PT useful for diagnosis, treatment and in screening for specific binding  
 PT agents.  
 XX  
 PS Claim 2; Page 1199; 1607pp; German.  
 XX  
 CC This invention describes novel cytostatic polynucleotide and polypeptide  
 CC sequences which can be used in a method for diagnosing prostatic cancer  
 CC or the risk of developing prostatic cancer. Diagnosis is based on  
 CC determining over transcription or over expression of the sequences in  
 CC prostatic tissue. Screening for inhibitors of the sequences or detection  
 CC substances involves a binding assay, any compounds that bind are  
 CC selected, optionally after deconvolution of mixtures. Detection of a  
 CC predetermined minimum level of the reporter indicates the presence of  
 CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,  
 CC short-interfering RNA or ribozymes; an organic molecule of molecular  
 CC weight below 5000, preferably 300, that binds to the polypeptide; an  
 CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the  
 CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human  
 CC (monoclonal) antibody directed against Ab or any of the above derivatised  
 CC with a reporter group, cell toxin, immunostimulatory molecules and/or  
 CC radioisotope. The polynucleotides are identified in human prostatic  
 CC cancer by differential expression analysis, using DNA microarrays,  
 CC between normal and tumorous tissues, with (over)expression being detected  
 CC by quantitative PCR. Analysis of prostatic cancer samples showed that  
 CC CD24 was upregulated in many of them. Sections of tissue, isolated from  
 CC prostatic cancer patients, or subjects at risk, were incubated;  
 CC sequentially with anti-human CD4 murine monoclonal antibodies;  
 CC biotinylated second antibody; streptavidin-conjugated horseradish  
 CC peroxidase and then diaminobenzidine as colour former (brown). The  
 CC samples were counterstained with hemalum (blue). Malignant cells stained  
 CC strongly but non-malignant cells only weakly. In 15 of 63 samples of  
 CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and  
 CC lymph node metastases were also stained. ADR65805-ADR65954 represent the  
 CC polynucleotide and polypeptide sequences used in the method of the  
 CC invention.  
 XX  
 XX Sequence 486 AA;  
 Query Match 100.0%; Score 48; DB 8; Length 486;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ILDDIGHGV 9  
 DB 156 ILDDIGHGV 164  
 RESULT 7  
 ADR66363  
 ID ADR66363 standard; protein; 486 AA.  
 XX  
 AC ADR66363;  
 XX  
 XX 02-DEC-2004 (first entry)  
 DT  
 XX Human prostatic carcinoma derived protein SEQ ID 217 #2.  
 DE  
 XX human; cytostatic; diagnosis; prostatic cancer;  
 KW



KW differential expression analysis.

XX Homo sapiens.

XX WO2004076614-A2.

XX 10-SEP-2004.

XX 22-FEB-2004; 2004WO-DE0000433.

XX 27-FEB-2003; 2003DE-01009985.

XX 14-MAY-2003; 2003DE-01022134.

XX (HINZ/) HINZMANN B.

XX (DAHL/) DAHL E.

XX (ROSE/) ROSENTHAL A.

XX (HERM/) HERMANN K.

XX (PILA/) PILARSKY C.

XX Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;

PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;

PI Xinzhang L, Staub E;

XX WPI; 2004-653386/63.

XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,  
PT useful for diagnosis, treatment and in screening for specific binding  
PT agents.

XX Claim 2; Page 697; 1607pp; German.

XX This invention describes novel cytostatic polynucleotide and polypeptide  
CC sequences which can be used in a method for diagnosing prostatic cancer  
CC or the risk of developing prostatic cancer. Diagnosis is based on  
CC determining over transcription or over expression of the sequences in  
CC prostatic tissue. Screening for inhibitors of the sequences or detection  
CC substances involves a binding assay, any compounds that bind are  
CC selected, optionally after deconvolution of mixtures. Detection of a  
CC predetermined minimum level of the reporter indicates the presence of  
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,  
CC short-interfering RNA or ribozymes; an organic molecule of molecular  
CC weight below 5000, preferably 300, that binds to the polypeptide; an  
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the  
CC polypeptide, preferably humanised or human; an anti-idiotype, non-human  
CC (monoclonal) antibody directed against Ab or any of the above derivatised  
CC with a reporter group, cell toxin, immunostimulatory molecules and/or  
CC radioisotope. The polynucleotides are identified in human prostatic  
CC cancer by differential expression analysis, using DNA microarrays,  
CC between normal and tumorous tissues, with (over)expression being detected  
CC by quantitative PCR. Analysis of prostatic cancer samples showed that  
CC CD24 was upregulated in many of them. Sections of tissue, isolated from  
CC prostatic cancer patients, or subjects at risk, were incubated  
CC sequentially with anti-human CD4 murine monoclonal antibodies;  
CC biotinylated second antibody; streptavidin-conjugated horseradish  
CC peroxidase and then diaminobenzidine as colour former (brown). The  
CC samples were counterstained with hemalum (blue). Malignant cells stained  
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of  
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and  
CC lymph node metastases were also stained. ADR65805-ADR66954 represent the  
CC polynucleotide and polypeptide sequences used in the method of the  
CC invention.

XX SQ Sequence 486 AA;

Query Match 100.0%; Score 48; DB 8; Length 486;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9

Db 156 ILDDIGHGV 164

RESULT 8

ADE63219

ID ADE63219 standard; protein; 543 AA.

XX AC ADE63219;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein AAC53493, SEQ ID NO 9156.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; AAC53493.

XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 543 AA;

Query Match 100.0%; Score 48; DB 7; Length 543;

Best Local Similarity 100.0%; Pred. No. 2.6;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9



XX AAW15775;  
 XX 31-OCT-1997 (first entry)  
 XX  
 XX Protein cognate of protein kinase C-theta (clone 2-18 product).  
 XX  
 XX Signal transduction; cell signalling; modulator; immunomodulator; human;  
 KW protein kinase C-theta; PKC-theta; cognate; graft rejection;  
 KW autoimmune disease; allergy; asthma; therapy.  
 XX  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 127 /note= "encoded by CVT"  
 FT Misc-difference 129 /note= "encoded by AVT"  
 FT Misc-difference 165 /note= "encoded by TGA"  
 FT Misc-difference 183 /note= "encoded by TAA"  
 FT  
 XX WO9714038-A1.  
 XX  
 XX 17-APR-1997.  
 PD  
 XX 10-OCT-1996; 96WO-US016195.  
 XX  
 XX 10-OCT-1995; 95US-00541964.  
 PR  
 XX 31-JAN-1996; 96US-00594447.  
 PR  
 XX 18-JUN-1996; 96US-00665647.  
 PR  
 XX (TERR-) TERRAPIN TECHNOLOGIES INC.  
 PA  
 XX Mochly-Rosen D, Ron D, Kauvar LM, Napolitano EW, Voronova A;  
 PI Vasquez NJ;  
 PI  
 XX WPI; 1997-236030/21.  
 DR  
 XX N-PSDB; AAT60590.  
 DR  
 XX Identifying a modulator of intracellular signal transduction - by  
 PT determining the interaction of a signal generating peptide with the test  
 PT substance, allows modulation of the immune system.  
 XX  
 XX Example 8; Fig 15C; 74pp; English.  
 PS  
 XX This sequence is a protein cognate of protein kinase C-theta (PKC- theta)  
 CC that shows homology with SH3-contg. C-abl binding proteins. The amino  
 CC acid sequence was deduced from cDNA clone 2-18 (AAT60590), isolated from  
 CC a human CD4+ T-cell cDNA library using a yeast two- hybrid system. A  
 CC claimed method for prep. of the cognate involves culturing recombinant  
 CC host cells under conditions where the cognate is produced. The cognate is  
 CC used in a claimed method for identifying modulators of intracellular  
 CC signal transduction. This involves assessing the ability of candidate  
 CC modulators to affect the interaction between a catalytically active  
 CC signal-generating protein (e.g. PKC-theta) and a cognate binding protein  
 CC involved in modulating the signal transduction function. Substances  
 CC identified using this method are used to modulate the immune system in a  
 CC subject (claimed). They act to reduce T-cell activity, reduce the rate of  
 CC graft rejection, reduce the severity of an autoimmune disorder,  
 CC ameliorate an allergic and/or asthmatic response, or diminish cytokine  
 CC response (claimed)  
 XX  
 XX Sequence 241 AA;  
 SQ  
 Query March 95.8%; Score 46; DB 2; Length 241;  
 Best Local Similarity 77.8%; Pred. No. 2.6;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ILDDIGHGV 9  
 :|||:||||  
 Db 101 VLDDVGHGV 109

RESULT 12  
 ADD71163  
 ID ADD71163 standard; protein; 367 AA.  
 XX  
 XX ADD71163;  
 AC  
 XX 15-JAN-2004 (first entry)  
 DT  
 XX Human intracellular signalling molecule INTSIG-52 protein SEQ ID NO:52.  
 DE  
 XX human; intracellular signalling molecule; INTSIG; cytostatic;  
 KW antiarteriosclerotic; antidiabetic; anticonvulsant; neurotropic;  
 KW neuroprotective; cerebroprotective; anti-HIV; antiallergic;  
 KW antiinflammatory; thymimetic; gene therapy;  
 KW cell proliferative disorder; cancer; atherosclerosis; endocrine disorder;  
 KW diabetes; neurological disorder; epilepsy; Huntington's disease; stroke;  
 KW immune disorder; inflammatory disorder; AIDS; allergy;  
 KW developmental disorder; Hypothyroidism; Cushing's syndrome;  
 KW reproductive disorder; vesicle-trafficking disorder; infection.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO2003039348-A2.  
 PN  
 XX 15-MAY-2003.  
 PD  
 XX 07-NOV-2002; 2002WO-US036151.  
 PF  
 XX 09-NOV-2001; 2001US-0344472P.  
 PR  
 XX 30-NOV-2001; 2001US-0334558P.  
 PR  
 XX 14-DEC-2001; 2001US-0340296P.  
 PR  
 XX 21-DEC-2001; 2001US-0343557P.  
 PR  
 XX 18-JAN-2002; 2002US-0350420P.  
 PR  
 XX 25-JAN-2002; 2002US-0351927P.  
 PR  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA  
 XX Kable AE, Swarnakar A, Gorvad AE, Hafalia AJA, Duggan BM;  
 PI Warren BA, Emerling BM, Ison CH, Nguyen DB, Lindquist EA, Lee EA;  
 PI Yue H, Yue H, Forsythe IJ, Rammuraj J, Griffin JA, Li JX;  
 PI Marquis JP, Gietzen KJ, Baughn MR, Borowsky ML, Yao MG, Chawla NK;  
 PI Lehn-Nason PM, Lal PG, Gururajan R, Khare R, Batra S, Becha SD;  
 PI Lee SY, Tran UK, Elliott VS, Sprague WW, Tang YT, Zebartadian Y;  
 PI Jiang X, Jackson AA, Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A;  
 PI Zheng W;  
 PI  
 XX WPI; 2003-441441/41.  
 DR  
 XX N-PSDB; ADD71215.  
 DR  
 XX New human intracellular signaling molecules (INTSIG)), useful for  
 PT diagnosing, treating and preventing diseases or conditions associated  
 PT with the aberrant INTSIG expression e.g. cancer, diabetes, epilepsy, or  
 PT infections.  
 PT  
 XX Claim 1; SEQ ID NO 52; 363pp; English.  
 PS  
 XX The present invention describes human intracellular signalling molecules  
 CC designated INTSIG-1 to INTSIG-52. The INTSIG polynucleotides and proteins  
 CC have cytostatic, antiarteriosclerotic, antidiabetic, anticonvulsant,  
 CC neurotropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic,  
 CC antiinflammatory and thymimetic activities, and can be used in gene  
 CC therapy. The INTSIG polynucleotides and proteins can be used in the  
 CC diagnosis, treatment and prevention of diseases or conditions associated  
 CC with the decreased expression or overexpression of INTSIG, such as cell  
 CC proliferative (e.g. cancer, atherosclerosis), endocrine (e.g. diabetes),  
 CC neurological (e.g. epilepsy, Huntington's disease, stroke),  
 CC immune/inflammatory (e.g. AIDS, allergies), developmental (e.g.  
 CC Hypothyroidism, Cushing's syndrome), reproductive and vesicle-trafficking  
 CC disorders, or infections. They are also useful in assessing the effects  
 CC of exogenous compounds on the expression of nucleic acid and amino acid  
 CC sequences of INTSIG. The INTSIG or its fragments are useful in screening

CC compounds for effectiveness as agonist or antagonist of the polypeptides,  
 CC or in altering the expression of the target polynucleotide and compounds  
 CC that specifically bind to or modulate the activity of the polypeptide.  
 CC The present sequence represents human INTSIG-52 from the present  
 CC invention.  
 XX  
 XX

SQ Sequence 367 AA;

Query Match 95.8%; Score 46; DB 7; Length 367;  
 Best Local Similarity 77.8%; Pred. No. 4;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9  
 :|||:||||  
 Db 145 VLDDVGHG 153

RESULT 13

ABB57113  
 ID ABB57113 standard; protein; 394 AA.

XX ABB57113;

XX 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related protein sequence SEQ ID NO:257.

XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease.

XX Mus musculus.

XX WO200188188-A2.

XX 22-NOV-2001.

XX 18-MAY-2001; 2001WO-JP004192.

XX 18-MAY-2000; 2000JP-00145977.

XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

DR WPI; 2002-034733/04.

DR N-PSDB; ABI99364.

XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
 PT expression levels of particular genes defined in the specification or by  
 PT determining the expression profile of a gene group comprising these  
 PT genes.

PS Claim 2; Page 726-728; 2690pp; English.

XX The present invention describes a method for examining ischaemic  
 CC conditions, comprising measuring the expression levels of particular  
 CC genes (I) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (I). The method  
 CC is useful for examining the ischaemic condition (e.g. compressive  
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding the  
 CC protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The expression  
 CC levels or expression profiles produced by these genes are used as an  
 CC indicator when screening for ischaemic condition-improving drugs or  
 CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR  
 CC primers for a mouse ischaemic condition related sequence, which are used  
 CC in the exemplification of the present invention

SQ Sequence 394 AA;

Query Match 95.8%; Score 46; DB 5; Length 394;  
 Best Local Similarity 77.8%; Pred. No. 4.3;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 ILDDIGHGV 9  
 :|||:||||  
 Db 60 VLDDVGHG 68

RESULT 14

ADR08679  
 ID ADR08679 standard; protein; 496 AA.

XX ADR08679;

XX 04-NOV-2004 (first entry)

XX Human protein useful for treating neurological disease Seq 2185.

XX human; oligo-capping method; diagnostic marker; gene therapy;  
 KW osteoporosis; neurological disease; Alzheimer's disease;  
 KW Parkinson's disease; dementia; short memory; cancer;  
 KW sense or motor function; emotional reaction; fear response; panic;  
 KW osteopathic; neuroprotective; neurotropic; antiparkinsonian; cytostatic;  
 KW tranquiliser.

XX Homo sapiens.

XX EPI447413-A2.

XX 18-AUG-2004.

XX 12-FEB-2004; 2004EP-00003145.

XX 14-FEB-2003; 2003JP-00102207.

XX 09-MAY-2003; 2003JP-00131452.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;  
 FI Wakamatsu A, Ishii S, Nagai K, Irie R;  
 XX WPI; 2004-583265/57.  
 DR N-PSDB; ADR06723.

XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,  
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
 XX Claim 1; SEQ ID NO 2185; 2686pp; English.

PS This invention relates to novel, isolated full length human cDNA

XX molecules and the encoded proteins thereof. Specifically, it refers to  
 CC cDNA clones obtained by an oligo-capping method, where none of these  
 CC clones are identical to any known human mRNAs. The present invention  
 CC describes an immunoassay to identify agonists and antagonists, as well as  
 CC antibodies, antisense molecules and siRNAs that can all be used to bind  
 CC to and modulate expression of the cDNA molecules. As such, these  
 CC molecules are useful for diagnostic markers or therapeutic targets for  
 CC the various diseases or morbid states. In particular, they are useful in  
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
 CC disease, Parkinson's disease, dementia, short memory and various cancers,  
 CC as well as for maintaining equilibrium of sense or motor function, and  
 CC for treating emotional reaction, fear response and panic. Accordingly,  
 CC they exhibit osteopathic, neuroprotective, neurotropic, antiparkinsonian,  
 CC cytoskeletal and tranquiliser activities. This polypeptide is a protein  
 CC encoded by a full length human cDNA sequence of the invention. NOTE: This  
 CC sequence is not given in the sequence listing of the specification but  
 CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-  
 CC office.

XX Sequence 496 AA;

Query Match 95.8%; Score 46; DB 8; Length 496;  
 Best Local Similarity 77.8%; Pred. No. 5.5;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILDDIGHGV 9  
 Db :|||:||||  
 162 VLDDVGHGV 170

RESULT 15  
 AAY83007  
 ID AAY83007 standard; protein; 365 AA.  
 XX AC AAY83007;  
 XX  
 XX  
 DT 04-JUL-2000 (first entry)  
 XX Human homer interacting protein I30.  
 DE  
 XX  
 KW Homer; calcium; receptor; immediate early gene; IEG; identification;  
 KW treatment; glutamate receptor; inositol triphosphate; epilepsy;  
 KW glutamate toxicity; memory disorder; learning disorder; stroke;  
 KW schizophrenia; Alzheimer's disease; tissue degeneration;  
 KW brain development; cardiac disorder; muscular disorder;  
 KW vascular disorder; neurological disorder; psychiatric disorder;  
 KW renal disorder; uterine disorder; bronchial disorder; ageing; human.  
 XX OS Homo sapiens.  
 XX  
 XX WO200011204-A2.  
 FN  
 XX  
 XX PD 02-MAR-2000.  
 XX  
 XX PF 18-AUG-1999; 99WO-US018973.  
 XX  
 XX PR 18-AUG-1998; 98US-0097334P.  
 PR 09-JUN-1999; 99US-0138426P.  
 PR 09-JUN-1999; 99US-0138493P.  
 PR 09-JUN-1999; 99US-0138494P.  
 XX  
 XX (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
 PA  
 XX  
 XX PI Worley PF, Tu JC, Xiao B, Leahy D, Beneken J, Lanahan AA;  
 XX  
 XX WPI; 2000-246571/21.  
 DR N-PSDB; RAZ93303.  
 XX  
 XX  
 PT Identifying compounds capable of modulating cellular response useful for  
 PT treating Alzheimer's disease and cardiac disorders, involves incubating  
 PT compound with cell expressing Homer protein and cell-surface receptor.  
 XX  
 PS Claim 65; Page 138; 171pp; English.  
 XX

XX Sequence 365 AA;  
 SQ  
 Query Match 89.6%; Score 43; DB 3; Length 365;  
 Best Local Similarity 87.5%; Pred. No. 14;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LDDIGHGV 9  
 Db 145 LDDIGHGI 152

Search completed: November 18, 2005, 01:01:15  
 Job time : 53.7742 secs

XX Homer proteins are the products of neuronal immediate early genes  
 (IEG's). They selectively bind the carboxy termini of certain cell-  
 surface receptors, certain intracellular receptors and binding proteins.  
 CC Many forms of Homer proteins contain a "coiled-coil" structure in the  
 CC carboxy terminal domain which mediated homo- and heteromultimerisation  
 CC between Homer proteins. Homer plays a significant role in mediating  
 CC receptor-activated calcium mobilisation from intracellular stores. Thus,  
 CC cells expressing a Homer protein can be used to identify a compound  
 CC capable of modulating a cellular response mediated by cell surface  
 CC receptor or intracellular receptor. Compounds identified in this manner  
 CC which modulate Homer protein activity are useful for treating disorders  
 CC associated with glutamate receptors such as epilepsy, glutamate toxicity,  
 CC memory disorders, disorders of learning, stroke, schizophrenia,  
 CC Alzheimer's disease, tissue degeneration and disorders of brain  
 CC development and also for treating disorders associated with Homer protein  
 CC activity which includes cardiac, muscular, vascular, neurological,  
 CC psychiatric, renal, uterine and bronchial tissue disorders and for  
 CC affecting the natural aging process. These compounds are also useful for  
 CC modulating receptor-mediated calcium mobilization, by exposing a cell to  
 CC the compound to modulate calcium mobilization that normally occurs when  
 CC the cell is exposed to a ligand, typically an agonist or antagonist of  
 CC metabotropic glutamate receptors, or to activate an intracellular  
 CC signaling pathway, especially an inositol triphosphate signaling pathway

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:11:06 ; Search time 8.90323 Seconds  
(without alignments)  
97.263 Million cell updates/sec

Title: US-10-006-177-7

Perfect score: 48

Sequence: 1 ILDDIGHGV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	390	2 G01936	Abl binding protei
2	40	83.3	436	2 T15331	hypothetical prote
3	38	79.2	209	2 E89009	protein R08E5.1 [i
4	37	77.1	409	2 C84307	hypothetical prote
5	36	75.0	331	2 G96750	unknown protein F2
6	36	75.0	344	2 C82835	anthranilate phosph
7	35	72.9	63	2 E84370	hypothetical prote
8	35	72.9	149	2 T38058	probable protein k
9	35	72.9	204	2 F84217	hypothetical prote
10	35	72.9	256	1 B83852	phosphoesterase-re
11	35	72.9	348	2 T35450	ABC transporter AT
12	35	72.9	393	2 B84400	hypothetical prote
13	35	72.9	399	2 E71188	hypothetical prote
14	35	72.9	403	2 G70038	conserved hypothet
15	35	72.9	405	2 B82871	conserved hypothet
16	35	72.9	417	2 E75148	hypothetical prote
17	35	72.9	432	2 B84127	hypothetical prote
18	35	72.9	471	2 T06287	hypothetical prote
19	35	72.9	1088	2 S39261	VPI protein - porc
20	35	72.9	1737	2 T00209	MEGF8 protein - hu
21	34	70.8	95	2 JC4321	gibberellin respon
22	34	70.8	100	2 A84462	hypothetical prote
23	34	70.8	180	2 H64342	hypothetical prote
24	34	70.8	186	1 F32354	transaldolase (pen
25	34	70.8	191	2 AE0412	probable 2'-5' RNA
26	34	70.8	212	2 A84123	transaldolase (pen
27	34	70.8	217	2 H64419	transaldolase (EC
28	34	70.8	376	2 S55149	hypothetical prote
29	34	70.8	379	2 JC5303	conserved hypothet

30	34	70.8	406	2 B69064	conserved hypothet
31	34	70.8	410	2 C83365	2-oxoisovalerate d
32	34	70.8	453	2 D86759	hypothetical prote
33	34	70.8	470	2 F72238	conserved hypothet
34	34	70.8	506	2 T04628	cytochrome P450 ho
35	34	70.8	535	2 T06285	hypothetical prote
36	34	70.8	587	2 T27765	hypothetical prote
37	34	70.8	668	2 S56909	polymyxin B resist
38	34	70.8	747	2 I39444	AMP deaminase (EC
39	34	70.8	839	2 T39190	probable ATP-depen
40	34	70.8	1242	2 G88480	protein C16A3.7 [i
41	33	68.8	127	2 D69186	hypothetical prote
42	33	68.8	135	2 H69202	hypothetical prote
43	33	68.8	140	2 A83335	fructose operon fucU
44	33	68.8	188	2 G81155	hypothetical prote
45	33	68.8	188	2 H81949	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

G01936

Abl binding protein 3 - human

C:Species: Homo sapiens (man)

C>Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 09-Jul-2004

C:Accession: G01936

R:Ren, R.

submitted to the EMBL Data Library, July 1995

A:Reference number: G08875

A:Accession: G01936

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-390 <REN>

A:Cross-references: UNIPROT:Q13249; EMBL:U31089; NID:G987264; PID:G987265

C:Genetics:

A:Gene: AblBP3

F:335-382/Domain: SH3 homology <SH3>

Query Match 100.0%; Score 48; DB 2; Length 390;  
Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9

Db 60 ILDDIGHGV 68

##### RESULT 2

T15331

hypothetical protein B0336.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T15331

R:Taich, A.

submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid B0336.

A:Reference number: Z18331

A:Accession: T15331

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-436 <TAI>

A:Cross-references: EMBL:U32305; NID:G912752; PID:G912758; PIDN:AAC46832.1; CESP:B0336.6

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:B0336.6

A:introns: 10/3; 39/2; 186/3; 384/1

Query Match 83.3%; Score 40; DB 2; Length 436;  
Best Local Similarity 77.8%; Pred. No. 7;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9

```
Db      114 VLDIGHGV 122
      :|| |||||
RESULT 3
E89009 protein R08E5.1 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: E89009
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: E89009
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-209 <STO>
A;Cross-references: UNIPROT:O01594; GB:chr_V; PIDN:AA852275.1; PID:gl938440; GSPDB:GN000
C;Genetics:
A;Gene: R08E5.1
A;Map position: 5

      Query Match      79.2%; Score 38; DB 2; Length 209;
      Best Local Similarity 77.8%; Pred. No. 7.4;
      Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ILDDIGHGV 9
      :|| |||||
Db      88 MLPDIGHGV 96

RESULT 4
C84307 hypothetical protein Vngi537c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: C84307
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabc
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-409 <STO>
A;Cross-references: UNIPROT:Q9HPP4; GB:AE004437; NID:gl0581025; PIDN:AAG19823.1; GSPDB:G
A;Gene: VNGI537C
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1154

      Query Match      77.1%; Score 37; DB 2; Length 409;
      Best Local Similarity 87.5%; Pred. No. 25;
      Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ILDDIGHG 8
      :|| |||||
Db      84 ILHDIGHG 91

RESULT 5
G96750 unknown protein F28P22.19 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: G96750
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Cressy, T.H.; Dewar, K.;
```

```
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Matti, R.; Marzfalli,
Rizzo, M.; Rooney, T.; Rowley, D.; Schwartz, J.R.; Sakano, H.
A;Authors: Salzberg, S.L.; Rowley, D.; Schwartz, J.R.; Venter, J.C.; Davis, R.W.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96750
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-331 <STO>
A;Cross-references: UNIPROT:Q9CAH7; GB:AE005173; NID:g6648166; PIDN:AAF21166.1; GSPDB:GN
C;Genetics:
A;Gene: F28P22.19
A;Map position: 1

      Query Match      75.0%; Score 36; DB 2; Length 331;
      Best Local Similarity 66.7%; Pred. No. 31;
      Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ILDDIGHGV 9
      :|| |||||
Db      303 VLKEIGHGV 311

RESULT 6
C82835 anthranilate phosphoribosyltransferase XF0212 [imported] - Xylella fastidiosa (strain 9a
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: C82835
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82835
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-344 <SIM>
A;Cross-references: UNIPROT:Q9PGT6; GB:AE003874; GB:AE003849; NID:g9105007; PIDN:AAF8302
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorfy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, M.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0212
C;Superfamily: anthranilate phosphoribosyltransferase; trpD homology

      Query Match      75.0%; Score 36; DB 2; Length 344;
      Best Local Similarity 66.7%; Pred. No. 32;
      Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ILDDIGHGV 9
      :|| |||||
Db      307 VAQDIGHGV 315

RESULT 7
E84370
```



```

SCOE:SCOE
GSPDB:GN00070;
FIDN:CAAL19904.1;
EMBL:AL0031035;
A/Cross-references: UNIPROT:O86751;
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SC6A9.19c
C/Superfamily: ATP-binding cassette homology

Query Match      72.9%; Score 35; DB 2; Length 348;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 ILDDIGHGV 9
          :|||: ||
Db      20 VLDDVGLGV 28

RESULT 12
B84400
hypothetical protein Vng2502c [imported] - Halobacterium sp. NRC-1

```

C;Species: Halobacterium sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: B84400  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabcic, Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A;Title: Genome sequence of Halobacterium species NRC-1.  
A;Reference number: A84160; MUID:20504483; PMID:11016950  
A;Accession: B84400  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-393 <STO>  
A;Cross-references: UNIPROT:Q9HMK5; GB:AE004437; NID:gl0581895; PIDN:AAG20566.1; GSPDB:C  
C;Genetics:  
A;Gene: VNG2502C  
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1154

Query Match 72.9%; Score 35; DB 2; Length 393;  
Best Local Similarity 75.0%; Pred. No. 58;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILDDIGHG 8  
|:|||||  
Db 83 LLHDIGHG 90

RESULT 13  
E71188  
hypothetical protein PH1782 - Pyrococcus horikoshii  
C;Species: Pyrococcus horikoshii  
C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 12-Jul-2004  
C;Accession: E71188  
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A;Reference number: A71000; MUID:98344137; PMID:9679194  
A;Accession: E71188  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-399 <KAW>  
A;Cross-references: UNIPROT:O59446; GB:AP000007; NID:g3236134; PIDN:BA030900.1; PID:g325  
A;Experimental source: strain O73  
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C;Genetics:  
A;Gene: PH1782

Query Match 72.9%; Score 35; DB 2; Length 399;  
Best Local Similarity 75.0%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILDDIGHG 8  
|:|||||  
Db 72 LLHDIGHG 79

RESULT 14  
G70038  
conserved hypothetical protein yvfP - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: G70038  
R;Kunst, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowaka, A.; Serof,  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: G70038  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-403 <KUN>  
A;Cross-references: UNIPROT:O07014; GB:Z99121; GB:AL009126; NID:g2635827; PIDN:CAB15416.  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: yvfP

Query Match 72.9%; Score 35; DB 2; Length 403;  
Best Local Similarity 66.7%; Pred. No. 60;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9  
|:|||||  
Db 223 ILDMVGHGI 231

RESULT 15  
B82871  
conserved hypothetical UUS98 [imported] - Ureaplasma urealyticum  
C;Species: Ureaplasma urealyticum  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 14-Sep-2001  
C;Accession: B82871  
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mi  
A;Reference number: A82870  
A;Accession: B82871  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-405 <GLA>  
A;Cross-references: GB:AE002158; GB:AF222894; NID:g6899599; PIDN:AAF31012.1; GSPDB:GN001  
A;Experimental source: serovar 3; biovar 1  
C;Genetics:  
A;Gene: UUS98  
A;Genetic code: SGC3  
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ1154

Query Match 72.9%; Score 35; DB 2; Length 405;  
Best Local Similarity 75.0%; Pred. No. 60;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILDDIGHG 8  
|:|||||  
Db 96 LLHDIGHG 103

Search completed: November 18, 2005, 01:09:38  
Job time : 12.0032 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:04:26 ; Search time 39.9677 Seconds  
(without alignments)  
115.311 Million cell updates/sec

Title: US-10-006-177-7

Perfect score: 48

Sequence: 1 ILDDIGHGV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	446	1 AB12 MOUSE	P62484 mus musculus
2	48	100.0	480	2 Q6AXH6	Q6axh6 mus musculus
3	48	100.0	487	2 Q6AXD2	Q6axd2 mus musculus
4	48	100.0	513	1 AB12 HUMAN	Q9nyb9 homo sapien
5	48	100.0	543	2 Q35823	Q35823 rattus norv
6	46	95.8	450	2 Q6AXD3	Q6axd3 mus musculus
7	46	95.8	476	1 AB11 RAT	Q9qzm5 rattus norv
8	46	95.8	476	2 Q68FB6	Q68fb6 xenopus tro
9	46	95.8	481	1 AB11 MOUSE	Q8cbw3 mus musculus
10	46	95.8	508	1 AB11 HUMAN	Q8lzp0 homo sapien
11	45	93.8	503	2 Q66IV2	Q66iv2 xenopus lae
12	44	91.7	286	2 Q8BZP5	Q8bzip5 mus musculus
13	44	91.7	367	1 AB13 MOUSE	Q8byz1 mus musculus
14	44	91.7	450	2 Q7Q2G7	Q7q2g7 anopheles g
15	44	91.7	503	2 Q8VY25	Q8vy25 arabidopsis
16	44	91.7	503	2 Q84WV5	Q84wv5 arabidopsis
17	43	89.6	366	1 AB13 HUMAN	Q9p2a4 homo sapien
18	43	89.6	367	2 Q6AYC6	Q6ayc6 rattus norv
19	40	83.3	469	2 Q10929	Q10929 caenorhabdi
20	40	83.3	473	2 Q7Y0S9	Q7y0s9 drosophila
21	39	81.2	938	2 Q66IL1	Q66il1 xenopus tro
22	39	81.2	1267	2 Q6FPR7	Q6fpr7 candida gla
23	38	79.2	218	2 Q01594	Q01594 caenorhabdi
24	38	79.2	355	1 QUEA_PHO1L	Q7n0i3 photorhabdu
25	38	79.2	528	2 Q6LRX8	Q6lrk8 photobacter
26	38	79.2	534	2 Q6P886	Q6p886 xenopus tro
27	37	77.1	163	2 Q8XIM9	Q8xim9 clostridium
28	37	77.1	353	2 Q7ABE2	Q7ab2 thermoplasm
29	37	77.1	409	2 Q9HPP4	Q9hpp4 halobacteri
30	37	77.1	410	2 Q9AJ85	Q9aj85 lactococcus
31	37	77.1	410	2 Q9XB0	Q9xb0 lactococcus

32	37	77.1	412	2	Q845D2	Q845d2 lactococcus
33	37	77.1	414	2	Q7DH38	Q7dh38 lactococcus
34	37	77.1	414	2	Q8GRC4	Q8grc4 lactococcus
35	37	77.1	414	2	Q9REY0	Q9rey0 lactococcus
36	37	77.1	534	2	Q6GPB8	Q6gpb8 xenopus lae
37	37	77.1	785	2	Q8TMK7	Q8tmk7 methanosaer
38	36	75.0	67	2	Q6ESN3	Q6esn3 oryza sativ
39	36	75.0	156	2	Q6YDB1	Q6ydb1 arachis hyp
40	36	75.0	156	2	Q6YDB4	Q6ydb4 arachis hyp
41	36	75.0	175	2	Q6H8F9	Q6h8f9 oryza sativ
42	36	75.0	176	2	Q8LK41	Q8lk41 capsicum an
43	36	75.0	209	2	Q8RT45	Q8rt45 nostoc linc
44	36	75.0	266	2	Q73AJ4	Q73aj4 bacillus ce
45	36	75.0	266	2	Q6HKP3	Q6hkp3 bacillus th

#### ALIGNMENTS

RESULT 1  
ABI2 MOUSE  
ID ABI2 MOUSE STANDARD; PRT; 446 AA.  
AC P62484; Q6PHU3;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DE Abi-interactor 2 (Abelson interactor 2) (Abi-2).  
GN Name=Abi2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.J., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
SUBCELLULAR LOCATION.  
PubMed=11516653; DOI=10.1016/S0960-9822(01)00239-1;  
RX Stradal T., Courtney K.D., Rottner K., Hahne P., Small J.V.,  
RA Pendergast A.M.;  
RT "The Abi interactor proteins localize to sites of actin polymerization  
at the tips of lamellipodia and filopodia.";  
RL Curr. Biol. 11:891-895(2001).  
[3]  
FUNCTION, AND SUBCELLULAR LOCATION.  
PubMed=15143189; DOI=10.1128/MCB.24.11.4979-4993.2004;  
RX Echarri A., Lai M.J., Robinson M.R., Pendergast A.M.;  
RA "Abi interactor 1 (Abi-1) wave-binding and SNARE domains regulate its  
nucleocytoplasmic shuttling, lamellipodium localization, and wave-1  
levels.";  
RL Mol. Cell. Biol. 24:4979-4993(2004).  
[4]

RP TISSUE SPECIFICITY. AND DEVELOPMENTAL STAGE.  
RX PubMed=1095551; DOI=10.1006/mcne.2000.0865;  
RA Courtney K.D., Grove M., Vandongen H., Vandongen A., LaMantia A.-S.,  
RA Pendergaat A.M.;  
RA "Localization and phosphorylation of Abl-interactor proteins, Abi-1  
RT and Abi-2, in the developing nervous system.";  
RL Mol. Cell. Neurosci. 16:244-257(2000).  
CC -1- FUNCTION: May act in regulation of cell growth and transformation  
CC by interacting with nonreceptor tyrosine kinases Abl1 and/or Abl2.  
CC May be involved in cytoskeletal reorganization. Regulates Abl1/c-  
CC Abl-mediated phosphorylation of MENA.  
CC -1- SUBUNIT: Interacts with Abl1 and Abl2.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Localized to protruding  
CC lamellipodia and filopodia tips.  
CC -1- TISSUE SPECIFICITY: Expresses in embryonic and adult brain. In  
CC adult brain prominently expressed in the neocortex, hippocampus  
CC and dentate gyrus.  
CC -1- DEVELOPMENTAL STAGE: Detected at E10 in developing brain, and  
CC expression is more prominent in the neuroepithelium compared to  
CC the surrounding tissue. At E12 expression is enhanced throughout  
CC the CNS and is detected along the full length of the spinal chord.  
CC At E16 expression remains enhanced in the CNS, and is particularly  
CC prominent in the olfactory bulb. Also highly expressed in dorsal  
CC root ganglia.  
CC -1- PTM: Is a substrate for Abl1 (By similarity).  
CC -1- SIMILARITY: Contains 1 SH3 domain.  
CC -1- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; BC056345; AAH56345.1; --  
DR MGD; MGI:106913; Abi2.  
DR PRODOM; PD000066; SH3; 1.  
DR PROSITE; PS50002; SH3; 1.  
DR PROSITE; PS50192; t-SNARE; 1.  
DR KW Coiled coil; Nuclear protein; Phosphorylation; SH3 domain.  
FT DOMAIN 45 107 t-SNARE coiled-coil homology.  
FT DOMAIN 384 443 SH3.  
FT DOMAIN 166 356 Pro-rich.  
SQ SEQUENCE 446 AA; 49387 MW; 8EC1FB543D93BEAA CRC64;

Query Match 100.0%; Score 48; DB 1; Length 446;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9  
Db 145 ILDDIGHGV 153

RESULT 2  
Q6AXH6 PRELIMINARY; PRT; 480 AA.  
AC Q6AXH6  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Abi2 protein.  
GN Name=Abi2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RA Director MGC Project;  
RA Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Contains 1 SH3 domain.  
DR EMBL; BC079545; AAH79545.1; --  
DR InterPro; IPR001018; Neu\_cyt\_fact\_2.  
DR InterPro; IPR01452; SH3.  
DR InterPro; IPR011511; SH3.2.  
DR InterPro; IPR000727; t-SNARE.  
DR Pfam; PF00018; SH3 1; 1.  
DR Pfam; PF07653; SH3 2; 1.  
DR PRINTS; PR00499; P67PHOX.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRODOM; PD000066; SH3; 1.  
DR SMART; SM00326; SH3; 1.  
DR PROSITE; PS50002; SH3; 1.  
DR PROSITE; PS50192; t-SNARE; 1.  
KW SH3 domain.  
SQ SEQUENCE 480 AA; 52840 MW; CB49EBF5E8541188 CRC64;

Query Match 100.0%; Score 48; DB 2; Length 480;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9  
Db 145 ILDDIGHGV 153

RESULT 3  
Q6AXD2 PRELIMINARY; PRT; 487 AA.  
AC Q6AXD2  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Abi2 protein.  
GN Name=Abi2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RA Director MGC Project;  
RA Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
DE EMBL: BC079646; AAH79646.1; -;  
DR InterPro: IPR001452; SH3.  
DR InterPro: IPR000727; T\_SNARE.  
DR Pfam: PF00018; SH3\_1; -;  
DR PRINTS: PR00452; SH3DOMAIN.  
DR PRODOM: PD000066; SH3; 1.  
DR PROSITE: PS00002; SH3; 1.  
DR PROSITE: PS0192; T\_SNARE; 1.  
KW SH3 domain.  
SQ SEQUENCE 487 AA; 52900 MW; 605CCE7FBADEAA6A CRC64;  
  
Query Match 100.0%; Score 48; DB 2; Length 487;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 ILDDIGHGV 9  
Db 145 ILDDIGHGV 153  
  
RESULT 4  
AB12 HUMAN  
ID AB12\_HUMAN STANDARD; PRT; 513 AA.  
AC Q9NYB9; Q13249; Q13801; Q9BV70;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE ABL-interactor 2 (Abelson interactor 2) (Abi-2) (Abl binding protein  
DE 3) (ABLBP3) (Arg binding protein 1) (ArgBP1).  
GN Name=AB12; Synonyms=ARGBP1A;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3), FUNCTION, PHOSPHORYLATION,  
RP SUBCELLULAR LOCATION, AND INTERACTION WITH ABL1.  
RX MEDLINE=96067151; PubMed=7590236;  
RA Dai Z., Pendergast A.M.;  
RT "Abi-2, a novel SH3-containing protein interacts with the c-Abl  
RT tyrosine kinase and modulates c-Abl transforming activity.";  
RL Genes Dev. 9:2569-2582 (1995).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC Ren R.;  
RT "Cloning of a binding substrate of the Abl protein tyrosine kinase.";  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, AND INTERACTION WITH ABL2.  
RC TISSUE=Brain;  
RX MEDLINE=96211399; PubMed=8649853;  
RA Wang B., Myalliwec T., Krainc D., Jensen R.A., Sonoda G., Testa J.R.,

RA Golemis E.A., Kruh G.D.;  
RT "Identification of ArgBP1, an Arg protein tyrosine kinase binding  
RT protein that is the human homologue of a CNS-specific Xenopus gene.";  
RL Oncogene 12:1921-1929 (1996).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RA Kainine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,  
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,  
RA Phelan M., Farmer A.;  
RT "Cloning of human full-length CDSs in BD Creator(TM) system donor  
RT vector.";  
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Placenta;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins P.S., Wagner L., Sherman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buétow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins L.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [6]  
RP FUNCTION.  
RX PubMed=10498863; DOI=10.1038/sj.onc.1202911;  
RA Juang J.L., Hoffmann F.M.;  
RT "Drosophila abelson interacting protein (dabi) is a positive regulator  
RT of abelson tyrosine kinase activity.";  
RL Oncogene 18:5138-5147 (1999).  
RN [7]  
RP SUBCELLULAR LOCATION.  
RX PubMed=11516653; DOI=10.1016/S0960-9822(01)00239-1;  
RA Stradal T., Courtney K.D., Rottner K., Hahne P., Small J.V.,  
RA Pendergast A.M.;  
RT "The Abl interactor proteins localize to sites of actin polymerization  
RT at the tips of lamellipodia and filopodia.";  
RL Curr. Biol. 11:891-895 (2001).  
CC -!- FUNCTION: May act in regulation of cell growth and transformation  
CC by interacting with nonreceptor tyrosine kinases ABL1 and/or ABL2.  
CC May be involved in cytoskeletal reorganization. Regulates ABL1/c-  
CC ABL-mediated phosphorylation of MENA.  
CC -!- SUBUNIT: Interacts with ABL1 and ABL2.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Isoform 1 but not isoform 3 is  
CC localized to protruding lamellipodia and filopodia tips (By  
CC similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=1; Synonyms=Abi-2b;  
CC IsoId=Q9NYB9-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q9NYB9-2; Sequence=VSP\_010761, VSP\_010762, VSP\_010763;  
CC Name=3; Synonyms=Abi-2a;  
CC IsoId=Q9NYB9-3; Sequence=VSP\_010759, VSP\_010760, VSP\_010761,  
CC VSP\_010762;  
CC -!- PTM: Is a substrate for ABL1.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC -!- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR ENBL; U23435; AAA92289.1; --  
 DR ENBL; U31089; AAA75446.1; --  
 DR ENBL; AF260261; AAF70308.1; --  
 DR ENBL; X95632; CAA64885.1; --  
 DR ENBL; BT009920; AAP88922.1; --  
 DR ENBL; BC001439; AAH01439.1; --  
 DR PIR; G01936; G01936.  
 DR HSP; P20929; IARK.  
 DR Genew; HGNC:24011; AB12.  
 DR MIM; 606442; --  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR00727; T\_SNARE.  
 DR Pfam; PF00018; SH3.1; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 DR PROSITE; PS0192; T\_SNARE; 1.  
 DR Alternative splicing; Coiled coil; Nuclear protein; Phosphorylation;  
 KW SH3 domain.  
 KW DOMAIN 45 107 t-SNARE coiled-coil homology.  
 FT DOMAIN 451 510 SH3.  
 FT DOMAIN 172 423 Pro-rich.  
 FT VARSPIC 1 45 Missing (in isoform 3).  
 FT VARSPIC 46 95 /FTid=VSP\_010759.  
 FT FT ALETKAYTQSLASVAYLINTLANNVLQMDIQASQLRRM  
 FT FT ESSINHSIQ -> MSCRCWISRHPSYEGWNLQSIIFKHQIR  
 FT FT GVDLESTFTVTKFGNCCSLRLNE (in isoform 3).  
 FT FT /FTid=VSP\_010760.  
 FT FT Missing (in isoform 2 and isoform 3).  
 FT FT /FTid=VSP\_010761.  
 FT FT Missing (in isoform 2 and isoform 3).  
 FT FT /FTid=VSP\_010762.  
 FT FT S -> SLAPPPPSILQVTPQLPLMGFVARQENIS (in  
 FT FT isoform 2).  
 FT FT /FTid=VSP\_010763.  
 FT FT S -> R (in Ref. 3).  
 FT FT A -> D (in Ref. 3).  
 FT FT S -> T (in Ref. 2).  
 FT FT G -> P (in Ref. 1).  
 FT FT A -> V (in Ref. 2).  
 FT FT F -> S (in Ref. 2).  
 FT FT  
 FT SEQUENCE 513 AA; 55663 MW; 822983A69E5EA512 CRC64;  
 Query Match 100.0%; Score 48; DB 1; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 ILDDIGHGV 9  
 Db 145 ILDDIGHGV 153  
 RESULT 5  
 ID O35823 PRELIMINARY; PRT; 543 AA.  
 AC O35823;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Thyroid hormone responsive protein.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=Pisher 344; TISSUE=Cerebral cortex;  
 RX MEDLINE=97465844; PubMed=9359437;  
 RA Shah G.N., Li J., Schneiderjohn P., Mooradian A.D.;  
 RT "Cloning and characterization of a complementary DNA for a thyroid  
 RT hormone-responsive protein in mature rat cerebral tissue.";  
 RL Biochem. J. 327:617-623(1997).  
 CC -1- SIMILARITY: Contains 1 SH3 domain.  
 DR ENBL; U94904; AAC53493.1; --  
 DR HSP; P20929; IARK.  
 DR InterPro; IPR00108; Neu\_cyt\_fact\_2.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF00018; SH3.1; 1.  
 DR PRINTS; PR00499; P67PHOX.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 KW SH3 domain.  
 KW SEQUENCE 543 AA; 60171 MW; 5B5B07D68FE5BBE2 CRC64;  
 Query Match 100.0%; Score 48; DB 2; Length 543;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 ILDDIGHGV 9  
 Db 75 ILDDIGHGV 83  
 RESULT 6  
 ID O6AXD3 PRELIMINARY; PRT; 450 AA.  
 AC O6AXD3;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalski U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Director MGC Project;  
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 SH3 domain.  
 DR ENBL; BC079642; AAH79642.1; --

DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001451; SH3\_2.  
 DR InterPro; IPR000727; T-SNARE.  
 DR Pfam; PF00018; SH3\_1; 1.  
 DR Pfam; PF07653; SH3\_2; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR PROSITE; PS50192; T-SNARE; 1.  
 KW Hypothetical protein; SH3 domain.  
 SQ SEQUENCE 450 AA; 48939 MW; 5ESD0530B08292E3 CRC64;  
 Query Match 95.8%; Score 46; DB 2; Length 450;  
 Best Local Similarity 77.8%; Pred. No. 2.9;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ILDDIGHGV 9  
 Db 145 VLDDVGHGV 153  
 RESULT 7  
 ABIL\_RAT STANDARD; PRT; 476 AA.  
 ID ABIL\_RAT  
 AC Q9QZM5.  
 DT 25-OCT-2004 (Ref. 45, Created)  
 DT 25-OCT-2004 (Ref. 45, Last sequence update)  
 DT Abi-interactor 1 (Abelson interactor 1) (Abi-1) (Eps8 SH3 domain  
 DE binding protein) (Eps8 binding protein) (e3b1).  
 GN Name=Abil;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP STRAIN=Sprague-Dawley;  
 RC Nicolaas G., Galand C., Lecomte M.C.;  
 RA "cDNA sequence of the rat eps8 binding protein (e3b1).";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SUBCELLULAR LOCATION.  
 RA PubMed=10995551; DOI=10.1006/mcne.2000.0865;  
 RA Courtney K.D., Grove M., Vandongen H., Vandongen A., LaMantia A.-S.,  
 RA Pendergast A.M.;  
 RA "Localization and phosphorylation of Abi-interactor proteins, Abi-1  
 and Abi-2, in the developing nervous system.";  
 RL Mol. Cell. Neurosci. 16:244-257(2000).  
 CC -!- FUNCTION: May act in negative regulation of cell growth and  
 CC transformation by interacting with nonreceptor tyrosine kinases  
 CC Abi1 and/or Abi2. May play a role in regulation of EGF-induced Erk  
 CC pathway activation. Involved in cytoskeletal reorganization and  
 CC EGRF signaling. Together with Eps8 participates in transduction of  
 CC signals from Ras to Rac. In vitro, a trimeric complex of Abi1,  
 CC Eps8 and SOS1 exhibits Rac specific guanine nucleotide exchange  
 CC factor (GEF) activity and Abi1 seems to act as an adapter in the  
 CC complex. Regulates Abi1/c-Abl-mediated phosphorylation of MENA.  
 CC Recruits WASF1 to lamellipodia and there seems to regulate WASF1  
 CC protein level (By similarity).  
 CC -!- SUBUNIT: Interacts with ABIL1, MENA, STX1A, SNAP25, VAMP2, EPS8,  
 CC SOS1, SOS2, GRB2, SPTA1, NAPI1, the first SH3 domain of NCK1 and  
 CC through its N-terminus with WASF1. Part of a complex consisting of  
 CC ABIL1, STX1A and SNAP25. Part of a complex consisting of ABIL1, EPS8  
 CC and SOS1.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).  
 CC Localized to protruding lamellipodia and filopodia tips. Also  
 CC localized to neuronal growth cones and synaptosomes.  
 CC -!- DOMAIN: The t-SNARE coiled-coil homology domain is necessary and  
 CC sufficient for interaction with STX1A (By similarity).  
 CC -!- PTM: Phosphorylated (By similarity). In vitro substrate for v-Abl.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.

CC -!- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; AF176784; AAD55263.1; -  
 DR HSSP; P20929; IARK.  
 DR RGD; 621008; Abil.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR000727; T-SNARE.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR PROSITE; PS50192; T-SNARE; 1.  
 KW Coiled coil; Cytoskeleton; Nuclear protein; Phosphorylation;  
 FT SH3 domain.  
 FT DOMAIN 45 107 t-SNARE coiled-coil homology.  
 FT DOMAIN 414 473 SH3.  
 FT DOMAIN 332 386 Pro-rich.  
 SQ SEQUENCE 476 AA; 51704 MW; A2B8CC377090BEA6 CRC64;  
 Query Match 95.8%; Score 46; DB 1; Length 476;  
 Best Local Similarity 77.8%; Pred. No. 3.1;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ILDDIGHGV 9  
 Db 145 VLDDVGHGV 153  
 RESULT 8  
 Q68FB6 PRELIMINARY; PRT; 476 AA.  
 ID Q68FB6  
 AC Q68FB6;  
 DT 25-OCT-2004 (TRENBLrel. 28, Created)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8364;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo.  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleorn M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";



RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBJ databases.  
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 SH3 domain.  
 DR ENBL; BC079926; AAH79926.1; "  
 DR InterPro; IPR00108; Neu\_Cyt\_fact\_2.  
 DR InterPro; IPR01452; SH3.  
 DR InterPro; IPR01151; SH3\_2.  
 DR InterPro; IPR00727; T\_SNARE.  
 DR Pfam; PF00018; SH3\_1; 1.  
 DR Pfam; PF07653; SH3\_2; 1.  
 DR PRINTS; PR00499; P67PHOX.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRODOM; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 DR PROSITE; PS50192; T\_SNARE; 1.  
 KW Hypothetical protein; SH3 domain.  
 SQ SEQUENCE 476 AA; 51548 MW; BEE7299902F09C1 CRC64;  
  
 Query Match 95.8%; Score 46; DB 2; Length 476;  
 Best Local Similarity 77.8%; Pred. No. 3.1;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 ILDPDIGHV 9  
 Db 145 VLDDVGHV 153  
 :|||:||||  
  
 RESULT 9  
 ID AB11 MOUSE STANDARD; PRT; 481 AA.  
 AC Q8CBW3; Q60747; Q912M5; Q92319; Q99KH4;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Ab1-interactor 1 (Abelson interactor 1) (Abi-1) (Spectrin SH3 domain  
 DE binding protein 1) (Eps8 SH3 domain binding protein) (Eps8 binding  
 DE protein) (e3b1) (Abphilin-1).  
 GN Name=Abi1; Synonyms=Ssh3bpl;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 4), ALTERNATIVE SPLICING (ISOFORM 2), AND  
 RP FUNCTION  
 RC STRAIN=BALB/c;  
 RX MEDLINE=21417640; PubMed=11526477; DOI=10.1038/sj/onc/1204502;  
 RA Ikeguchi A., Yang H.-Y., Gao G., Goff S.P.;  
 RT "Inhibition of v-Abl transformation in 3T3 cells overexpressing  
 RT different forms of the Abelson interactor protein Abi-1.";  
 RL Oncogene 20:4926-4934 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX PubMed=11477655; DOI=10.1002/gcc.1160;  
 RA Shibuya N., Taki T., Mugishima H., Chin M., Tsuchida M., Sako M.,  
 RA Kawa K., Ishii E., Miura I., Yanagisawa M., Hayaishi Y.;  
 RT "t(10;11)-acute leukemias with MLL-AF10 and MLL-AB11 chimeric  
 RT transcripts: specific expression patterns of AB11 gene in leukemia and  
 RT solid tumor cell lines.";  
 RL Genes Chromosomes Cancer 32:1-10 (2001).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikiado I., Osato N., Saito K., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Clothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierzki R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran M.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yaunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haie F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
 RA Fahey J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,  
 RA Whiting M., Madan A., Young A.W., Green E.D., Dickson M.C.,  
 RA Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [5]  
 RP SEQUENCE OF 70-360 FROM N.A. (ISOFORM 4), FUNCTION, PHOSPHORYLATION,  
 RP TISSUE SPECIFICITY AND INTERACTIONS WITH ABL1 AND V-ABL.  
 RX MEDLINE=96067152; PubMed=7590237;  
 RA Shi Y., Alin K., Goff S.P.;  
 RT "Abl-interactor-1, a novel SH3 protein binding to the carboxy-terminal  
 RT portion of the Abl protein, suppresses v-abl transforming activity.";  
 RL Genes Dev. 9:2583-2597 (1995).  
 RN [6]  
 RP FUNCTION, AND IDENTIFICATION IN A COMPLEX WITH EPS8 AND SOS1.  
 RX PubMed=10499589; DOI=10.1038/45822;  
 RA Scita G., Nordstrom J., Carbone R., Tenca P., Giardina G., Gutkind S.,  
 RA Bjarnegard M., Betscholtz C., Di Fiore P.P.;  
 RT "EPS8 and E3B1 transduce signals from Ras to Rac.";  
 RL Nature 401:290-293 (1999).  
 RN [7]  
 RP FUNCTION, AND INTERACTION WITH MENA.  
 RX PubMed=12672821; DOI=10.1074/jbc.M301447200;  
 RA Tani K., Sato S., Sukezane T., Kojima H., Hirose H., Hanafusa H.,  
 RA Shishido T.;  
 RT "Abl-interactor 1 promotes tyrosine 296 phosphorylation of mammalian  
 RT enabled (Mena) by c-Abl kinase.";  
 RL J. Biol. Chem. 278:21685-21692 (2003).  
 RN [8]  
 RP FUNCTION, SUBCELLULAR LOCATION, AND INTERACTIONS WITH STX1A; SNAP25;



RP VAMP2 AND WASF1.  
RX PubMed=15143189; DOI=10.1128/MCB.24.11.4979-4993.2004;  
RA Echarri A., Lei M.J., Robinson M.R., Pendergast A.M.;  
RT "Abl interactor 1 (Abl-1) wave-binding and SNARE domains regulate its  
RT nucleocytoplasmic shuttling, lamellipodium localization, and wave-1  
RT levels.";  
RL Mol. Cell. Biol. 24:4979-4993(2004).  
RN [9]  
RP TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.  
RX PubMed=10995551; DOI=10.1006/mcne.2000.0865;  
RA Courtney K.D., Grove M., Vandongen H., Vandongen A., LaMantia A.-S.,  
RA Pendergast A.M.;  
RT "Localization and phosphorylation of Abl-interactor proteins, Abl-1  
RT and Abl-2, in the developing nervous system.";  
RL Mol. Cell. Neurosci. 16:244-257(2000).  
RN [10]  
RP SUBCELLULAR LOCATION.  
RX PubMed=1151653; DOI=10.1016/S0960-9822(01)00239-1;  
RA Stradal T., Courtney K.D., Rottner K., Hahne P., Small J.V.,  
RA Pendergast A.M.;  
RT "The Abl interactor proteins localize to sites of actin polymerization  
RT at the tips of lamellipodia and filopodia.";  
RL Curr. Biol. 11:891-895(2001).  
RN [1]  
CC -!- FUNCTION: May act in negative regulation of cell growth and  
CC transformation by interacting with nonreceptor tyrosine kinases  
CC Abl1 and/or Abl2. In vitro, at least isoform 2 and isoform 4  
CC suppress the transforming activity of Abelson murine leukemia  
CC virus (v-Abl) after overexpression in fibroblasts. May play a role  
CC in regulation EGF-induced Erk pathway activation. Involved in  
CC cytoskeletal reorganization and EGFR signaling. Together with EPS8  
CC participates in transduction of signals from Ras to Rac. In vitro,  
CC a trimeric complex of Abl1, EPS8 and SOS1 exhibits Rac specific  
CC guanine nucleotide exchange factor (GEF) activity and Abl1 seems  
CC to act as an adapter in the complex. Regulates Abl1/c-Abl-mediated  
CC phosphorylation of MENA (By similarity). Recruits WASF1 to  
CC lamellipodia and there seems to regulate WASF1 protein level.  
CC -!- SUBUNIT: Interacts with MENA, Abelson murine leukemia virus v-Abl,  
CC Abl1, STX1A, SNAP25, VAMP2, and through its N-terminus with WASF1.  
CC Part of a complex consisting of Abl1, STX1A and SNAP25. Part of a  
CC complex consisting of Abl1, EPS8 and SOS1. Interacts with EPS8,  
CC SOS1, SOS2, GRB2, SPTA1, NAP1 and the first SH3 domain of NCK1 (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear. Localized to  
CC protruding lamellipodia and filopodia tips. Also localized to  
CC neuronal growth cones and synaptosomes (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=1;  
CC IsoId=O8CBW3-1; Sequence=Displayed;  
CC Name=2; Synonyms=Short;  
CC IsoId=O8CBW3-2; Sequence=VSP\_010756, VSP\_010757, VSP\_010758;  
CC Name=3;  
CC IsoId=O8CBW3-3; Sequence=VSP\_010756;  
CC Name=4; Synonyms=long;  
CC IsoId=O8CBW3-4; Sequence=VSP\_010756, VSP\_010757;  
CC -!- TISSUE SPECIFICITY: Widely expressed with highest levels in bone  
CC marrow, spleen, brain, testes, and embryonic brain. In adult brain  
CC prominently expressed in the neocortex, hippocampus and dentate  
CC gyrus.  
CC -!- DEVELOPMENTAL STAGE: Detected at E10 and E12 in developing brain,  
CC but does not appear more prominent in the neuroepithelium compared  
CC to the surrounding tissue.  
CC -!- DOMAIN: The t-SNARE coiled-coil homology domain is necessary and  
CC sufficient for interaction with STX1A.  
CC -!- PTM: In vitro substrate for v-Abl. Phosphorylated on tyrosine  
CC residues after serum stimulation or induction by v-Abl (By  
CC similarity).  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC -!- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF420251; AAL16036.1; -;  
DR EMBL; AY033645; AAX59381.1; -;  
DR EMBL; AK034476; BAC28722.1; -;  
DR EMBL; BC004657; AAO04657.1; -;  
DR EMBL; U17698; AAR00373.1; ALT\_SEQ.  
DR HSSP; P20929; IARK.  
DR IntAct; Q8CBW3; -;  
DR MGD; MGI:104913; Abil.  
DR GO; GO:0030027; C:lamellipodium; IDA.  
DR GO; GO:0005515; F:protein binding; IPI.  
Query Match 95.8%; Score 46; DB 1; Length 481;  
Best Local Similarity 77.8%; Pred. No. 3.1;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ILDDIGHGV 9  
Db :||:||||  
145 VLDDVGHGV 153  
RESULT 10  
ABIL HUMAN  
ID ABIL HUMAN STANDARD; PRT; 508 AA.  
AC Q81ZP0; O15147; Q76049; Q95060; Q8TB63; Q96S81; Q9NXZ9; Q9NYB8;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Abl-interactor 1 (Abelson interactor 1) (Abl-1) (Spectrin SH3 domain  
DE binding protein 1) (Eps8 SH3 domain binding protein) (Eps8 binding  
DE protein) (e3B1) (Nap1 binding protein) (Nap1BP) (Abl binding protein  
DE 4) (AblBP4)  
GN Name=ABIL; Synonyms=SSH3BP1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 6), PHOSPHORYLATION, SUBCELLULAR LOCATION,  
RP AND INTERACTIONS WITH EPS8 AND ABL1.  
RX MEDLINE=97163405; PubMed=9010225; DOI=10.1038/sj.onc.1200822;  
RA Biesova Z., Piccoli C., Wong W.T.;  
RT "Isolation and characterization of e3B1, an eps8 binding protein that  
RT regulates cell growth.";  
RL Oncogene 14:233-241(1997).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1), ALTERNATIVE SPLICING (ISOFORMS 3; 7; 8  
RP AND 9), AND INTERACTION WITH SPTA1.  
RX MEDLINE=98256293; PubMed=9593709; DOI=10.1074/jbc.273.22.13681;  
RA Ziennicka-Kotula D., Xu J., Gu H., Potempska A., Kim K.S.,  
RA Jenkins E.C., Trenkner E., Kotula L.;  
RT "Identification of a candidate human spectrin Src homology 3 domain-  
RT binding protein suggests a general mechanism of association of  
RT tyrosine kinases with the spectrin-based membrane skeleton.";  
RL J. Biol. Chem. 273:13681-13692(1998).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 4).  
RC TISSUE=T-cell;  
RA Wilson L.A., Fields D., Cruz L., Friesen J., Siminovich K.A.;  
RT "A new member of the Abl interactor protein family, AblBP4.";  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RA Quackenbush R.C., Pendergast A.M.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 3), TISSUE SPECIFICITY, AND INTERACTIONS  
RP WITH ABL1; NAP1 AND NCK1.  
RX MEDLINE=21311631; PubMed=11418237; DOI=10.1016/S0378-1119(01)00521-2;

RA Yamanoto A., Suzuki T., Sakaki Y.;  
RT "Isolation of hNap1BP which interacts with human Nap 1 (NCKAP1) whose  
RL expression is down-regulated in Alzheimer's disease.";  
RL Gene 271:159-169(2001).  
[6]  
RN SEQUENCE FROM N.A. (ISOFORM 5), AND INTERACTION WITH NCK1.  
RP TISSUE=Umbilical vein endothelial cells;  
RC PubMed=12681507; DOI=10.1016/S0014-5793(03)00262-X;  
RX Gu Y., Souza R.F., Wu R.F., Xu Y.C., Terada L.S.;  
RA "Induction of colonic epithelial cell apoptosis by p47-dependent  
RT oxidants(1).";  
RL FEBS Lett. 540:195-200(2003).  
[7]  
RN SEQUENCE FROM N.A.  
RP Chikri M.M., Boutin M.P., Vaxillaire M.M., Froguel M.P.;  
RA "In silico cloning of the human SSH3BP1/e3B1 gene.";  
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[8]  
RN SEQUENCE FROM N.A. (ISOFORM 5).  
RP TISSUE=Muscle;  
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,  
RA Bha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Roak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krywinski M.I., Skalska U., Smalish D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[9]  
RN FUNCTION, AND INTERACTIONS WITH SOS1, SOS2 AND GRB2.  
RX PubMed=11003655; DOI=10.1128/MCB.20.20.7591-7601.2000;  
RA Fan P.-D., Goff S.P.;  
RT "Abl interactor 1 binds to sos and inhibits epidermal growth factor-  
RT and v-Abl-induced activation of extracellular signal-regulated  
RT kinases";  
RL Mol. Cell. Biol. 20:7591-7601(2000).  
[10]  
RN ALTERNATIVE SPLICING (ISOFORMS 2 AND 10), AND CHROMOSOMAL  
RP TRANSLOCATION.  
RX PubMed=9694699;  
RA Taki T., Shibuya N., Taniwaki M., Hanada R., Morishita K., Bessho F.,  
RA Yanagisawa M., Hayashi Y.;  
RT "Abl-1, a human homolog to mouse Abl-interactor 1, fuses the MLL gene  
RL in acute myeloid leukemia with t(10;11)(p11.2;q23).";  
RL Blood 92:1125-1130(1998).  
CC -!- FUNCTION: May act in negative regulation of cell growth and  
CC transformation by interacting with nonreceptor tyrosine kinases  
CC Abl and/or ABL2. May play a role in regulation of EGF-induced Erk  
CC pathway activation. Involved in cytoskeletal reorganization and  
CC EGFR signaling. Together with Eph8 participates in transduction of  
CC signals from Ras to Rac. In vitro, a trimeric complex of Abl1,  
CC Eph8 and SOS1 exhibits Rac specific guanine nucleotide exchange  
CC factor (GEF) activity and Abl1 seems to act as an adapter in the  
CC complex. Regulates Abl1/c-Abl-mediated phosphorylation of MENA.  
CC Recruits WASF1 to lamellipodia and there seems to regulate WASF1  
CC protein level (By similarity).  
CC -!- SUBUNIT: Interacts with Abl1, MENA, STX1A, SNAP25, VAMP2, Eph8,  
CC and through its N-terminus with WASF1. Part of a complex  
CC consisting of Abl1, STX1A and SNAP25. Part of a complex consisting  
CC of Abl1, Eph8 and SOS1 (By similarity). Interacts with SOS1, SOS2,  
CC

CC GRB2, SPTA1, NAP1 and the first SH3 domain of NCK1. Isoform 6 does  
CC not interact with NCK1.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear. Localized to  
CC protruding lamellipodia and filopodia tips. Also localized to  
CC neuronal growth cones and synaptosomes (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=10;  
CC Comment=Additional isoforms seem to exist;  
CC Name=1;  
CC IsoId=Q8IZP0-1; Sequence=Displayed;  
CC Name=2; Synonyms=long, B48;  
CC IsoId=Q8IZP0-2; Sequence=VSP\_010749, VSP\_010750, VSP\_010751,  
CC VSP\_010752;  
CC Name=3;  
CC IsoId=Q8IZP0-3; Sequence=VSP\_010750, VSP\_010752;  
CC Name=4;  
CC IsoId=Q8IZP0-4; Sequence=VSP\_010750, VSP\_010751, VSP\_010752;  
CC Name=5;  
CC IsoId=Q8IZP0-5; Sequence=VSP\_010749, VSP\_010750;  
CC Name=6;  
CC IsoId=Q8IZP0-6; Sequence=VSP\_010750, VSP\_010751;  
CC Name=7; Synonyms=4;  
CC IsoId=Q8IZP0-7; Sequence=VSP\_010750, VSP\_010751, VSP\_010754,  
CC VSP\_010755;  
CC Name=8; Synonyms=5;  
CC IsoId=Q8IZP0-8; Sequence=VSP\_010750, VSP\_010751, VSP\_010752,  
CC VSP\_010754;  
CC Name=9; Synonyms=2;  
CC IsoId=Q8IZP0-9; Sequence=VSP\_010750;  
CC Name=10; Synonyms=B30;  
CC IsoId=Q8IZP0-10; Sequence=VSP\_010749, VSP\_010750, VSP\_010751,  
CC VSP\_010752, VSP\_010753, VSP\_010754;  
CC -!- TISSUE SPECIFICITY: Widely expressed with highest expression in  
CC BRAIN.  
CC -!- DOMAIN: The t-SNARE coiled-coil homology domain is necessary and  
CC sufficient for interaction with STX1A (By similarity).  
CC -!- PTM: In vitro substrate for v-Abl (By similarity). Phosphorylated  
CC on tyrosine residues after serum stimulation or induction by v-  
CC Abl.  
CC -!- DISEASE: Involved in acute leukemias through a chromosomal  
CC translocation t(10;11)(p11.2;q23) that involves Abl1 and MLL. Abl1  
CC isoform 2 was found to be present in acute leukemia MLL-Abl1  
CC fusion transcript.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC -!- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; AF006516; AAC62569.1; -  
CC EMBL; U87166; AAC39757.1; -  
CC EMBL; AF001628; AAC00897.1; -  
CC EMBL; AF260282; AAF70309.1; -  
CC EMBL; AB040151; BAB55675.1; -  
CC EMBL; AF540955; AAN28379.1; -  
CC EMBL; AJ277065; CAB88006.1; -  
CC EMBL; AJ277066; CAB88006.1; JOINED.  
CC EMBL; AJ277067; CAB88006.1; JOINED.  
CC EMBL; AJ277068; CAB88006.1; JOINED.  
CC EMBL; AJ277069; CAB88006.1; JOINED.  
CC EMBL; AJ277070; CAB88006.1; JOINED.  
CC EMBL; AJ277071; CAB88006.1; JOINED.  
CC EMBL; AJ277072; CAB88006.1; JOINED.  
CC EMBL; AJ277073; CAB88006.1; JOINED.  
CC EMBL; AJ277074; CAB88006.1; JOINED.  
CC EMBL; BC024254; AAH24254.1; -  
CC HSSP; P20929; IARK.  
CC IntAct; Q8IZP0; -.

DR Genew; HGNC:11320; ABI1.  
 DR MIM; 603050; C:cytosol; TAS.  
 DR GO; GO:0005829; C:endoplasmic reticulum; TAS.  
 DR GO; GO:0005783; C:soluble fraction; TAS.  
 DR GO; GO:0005635; C:soluble fraction; TAS.  
 DR GO; GO:0008092; P:cytoskeletal protein binding; TAS.  
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; TAS.

Query Match 95.8%; Score 46; DB 1; Length 508;  
 Best Local Similarity 77.8%; Pred. No. 3.3;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILDDIGHGV 9  
 Db 145 VLDDVGHGV 153  
 :|||||:

RESULT 11  
 Q66IV2 PRELIMINARY; PRT; 503 AA.  
 AC Q66IV2  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE MGC84356 protein.  
 GN Name=MGC84356;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Eye;  
 RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Eye;  
 RC PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Whiting J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakesley R.W., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Eye;  
 RC Klein S., Gerhard D.S.;  
 RA Submitted (AUG-2004) to the EMBL/GenBank/DBSJ databases.  
 CC -I- SIMILARITY: Contains 1 SH3 domain.  
 DR EMBL; BC081178; AAH81178.1; -  
 DR InterPro; IPR001018; Neu\_cyt\_fact\_2.

DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR011511; SH3 2.  
 DR Pfam; PF00018; SH3 1; 1.  
 DR Pfam; PF07653; SH3 2; 1.  
 DR PRINTS; PR00499; P67PHOX.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRODOM; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR PROSITE; PS50192; T\_SNARE; 1.  
 KW SH3 domain.  
 SQ SEQUENCE 503 AA; 54100 MW; DED356AE564A2C09 CRC64;

Query Match 93.8%; Score 45; DB 2; Length 503;  
 Best Local Similarity 66.7%; Pred. No. 5;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILDDIGHGV 9  
 Db 145 VLDDVGHGV 153  
 :|||||:

RESULT 12  
 Q8BZP5 PRELIMINARY; PRT; 286 AA.  
 ID Q8BZP5  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched  
 DE library, clone:930129K12 product:similar to NESH (NESH PROTEIN).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RC The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 [5]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,

```

RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hangaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK033965; BAC28531.1; -.
SQ SEQUENCE 286 AA; 29701 MW; B920E72449586DC3 CRC64;

Query Match 91.7%; Score 44; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LDDIGHGV 9
Db 149 LDDIGHGV 156
|||||||

RESULT 13
AB13_MOUSE STANDARD; PRT; 367 AA.
AC Q8BYZ1; Q8PE63; Q9D7S4;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE ABI gene family member 3 (New molecule including SH3) (Nesh).
GN Name=Ab13; Synonym=Nesh;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=C57BL/6J; TISSUE=Skin, and Stomach;
RX PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Niikado I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schombach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Guscinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sample C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang N.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,

```

```

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Bork S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP FUNCTION.
RX PubMed=11956071;
RA Ichigotani Y., Yokozaki S., Fukuda Y., Hamaguchi M., Matsuda S.;
RT "Forced expression of NESH suppresses motility and metastatic
RT dissemination of malignant cells.";
RL Cancer Res. 62:2215-2219(2002).
CC -!- FUNCTION: Inhibits ectopic tumor cell metastasis of SRD cells. In
CC vitro, reduces cell motility.
CC -!- SUBUNIT: May interact with PAK1 and PAK2. Probably interacts with
CC TARSH (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8BYZ1-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8BYZ1-2; Sequence=VSP_010773;
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AK008928; BAB25972.1; -
CC EMBL: AK037144; BAC29720.1; -
CC EMBL: BC058260; AAH58260.1; -
CC HSSP: P29354; ICQ.
CC MGD: MGI:1913860; Abi3.
CC ProDom: PD000066; SH3; 1.
CC PROSITE: PS50002; SH3; 1.
CC Alternative splicing; Coiled coil; SH3 domain.
CC DOMAIN 309 367
CC DOMAIN 36 64
CC DOMAIN 235 307
CC VARSPLFC 1 53
CC /FTid=VSP_010773.

```

```

FT CONFLICT 339 339 I -> T (in Ref. 1).
SQ SEQUENCE 367 AA; 39118 MW; 0CEC0279B8BAFF7 CRC64;

Query Match
Best Local Similarity 91.7%; Score 44; DB 1; Length 367;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDDIGHGV 9
DB 149 LDDIGHGV 156

RESULT 14
Q702G7 PRELIMINARY; PRT; 450 AA.
AC Q702G7;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP6813 (Fragment).
GN Name=agCG52894; ORFNames=ENSANGG00000017598;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008973; EAA13429.1; -.
DR HSSP; Q9NZM3; IUDL.
DR InterPro; IPR00108; Neu_cyt_fact_2.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3_1; 1.
DR PRINTS; PR00499; P67PHOX.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 1.
DR PROSITE; PS00002; SH3; 1.
KW SH3 domain.
FT NON_TER 1
SQ SEQUENCE 450 AA; 50100 MW; B20975EBE8545104 CRC64;

Query Match
Best Local Similarity 66.7%; Score 44; DB 2; Length 450;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILDDIGHGV 9
DB 152 LLDDVGHGI 160

RESULT 15
Q8VY25 PRELIMINARY; PRT; 503 AA.
AC Q8VY25;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein At2g44970.
GN Name=At2g44970;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,

```

```

RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074267; AAL66964.1; -.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 503 AA; 54739 MW; BFF7D9CC9A214292 CRC64;

Query Match
Best Local Similarity 91.7%; Score 44; DB 2; Length 503;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILDDIGHGV 9
DB 196 ILEDIGHGV 204

Search completed: November 18, 2005, 01:08:14
Job time : 43.9677 secs

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 00:48:52 ; Search time 13.6452 Seconds  
(without alignments)  
49.237 Million cell updates/sec

Title: US-10-006-177-7

Perfect score: 48

Sequence: 1 ILDDIGHGV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS-COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	47	2	US-08-549-004A-13
2	48	100.0	47	3	US-09-051-982A-13
3	48	100.0	401	2	US-08-549-004A-5
4	48	100.0	401	3	US-09-051-982A-5
5	46	95.8	239	2	US-08-665-647-11
6	46	95.8	480	3	US-09-189-035-5
7	46	95.8	480	3	US-09-382-086-5
8	46	95.8	531	4	US-09-949-016-8074
9	43	89.6	366	4	US-09-377-285B-20
10	43	89.6	370	4	US-09-377-285B-16
11	43	89.6	476	3	US-09-189-035-1
12	43	89.6	476	3	US-09-382-086-1
13	40	83.3	552	4	US-09-252-991A-27032
14	38	79.2	362	4	US-09-252-991A-20079
15	38	79.2	483	4	US-09-252-991A-23024
16	36	75.0	255	4	US-09-489-039A-9075
17	36	75.0	299	4	US-09-543-681A-5982
18	36	75.0	465	4	US-09-252-991A-29387
19	36	75.0	677	4	US-09-252-991A-20406
20	35	72.9	88	4	US-09-727-739B-3
21	35	72.9	114	4	US-09-727-739B-3
22	35	72.9	217	4	US-09-252-991A-22705
23	35	72.9	321	4	US-09-252-991A-24233
24	35	72.9	585	4	US-09-252-991A-32321
25	34	70.8	184	3	US-09-384-162-12
26	34	70.8	216	1	US-08-546-712-2
27	34	70.8	216	2	US-08-751-105-2

28	34	70.8	267	2	US-07-857-224B-37	Sequence 37, Appl
29	34	70.8	317	4	US-09-134-000C-5669	Sequence 5669, Ap
30	34	70.8	367	3	US-08-888-429A-13	Sequence 13, Appl
31	34	70.8	367	4	US-09-593-653-13	Sequence 13, Appl
32	34	70.8	376	4	US-10-081-644-4	Sequence 4, Appl
33	34	70.8	385	4	US-09-252-991A-21405	Sequence 21405, A
34	34	70.8	387	4	US-09-252-991A-22990	Sequence 22990, A
35	34	70.8	464	4	US-09-252-991A-16660	Sequence 16660, A
36	34	70.8	514	4	US-09-252-991A-30600	Sequence 30600, A
37	34	70.8	580	4	US-09-252-991A-24206	Sequence 24206, A
38	34	70.8	668	1	US-08-530-950-13	Sequence 13, Appl
39	34	70.8	668	3	US-09-149-879-13	Sequence 13, Appl
40	34	70.8	668	4	US-09-057-009-13	Sequence 13, Appl
41	33	68.8	62	4	US-09-252-991A-23533	Sequence 23533, A
42	33	68.8	224	4	US-09-583-110-2737	Sequence 2737, Ap
43	33	68.8	245	4	US-09-252-991A-16937	Sequence 16937, A
44	33	68.8	250	4	US-09-291-170A-12	Sequence 12, Appl
45	33	68.8	250	4	US-09-724-884-12	Sequence 12, Appl

#### ALIGNMENTS

##### RESULT 1

US-08-549-004A-13  
; Sequence 13, Application US/08549004A  
; Patent No. 5969101  
; GENERAL INFORMATION:  
; APPLICANT: PENDERGAST, ANN MARIE  
; APPLICANT: DAI ZONGHAN  
; TITLE OF INVENTION: ABL-INTERACTOR PROTEIN  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/549,004A  
; FILING DATE: 27-OCT-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-103  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 47 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-549-004A-13

Query Match 100.0%; Score 48; DB 2; Length 47;  
Best Local Similarity 100.0%; Pred. No. 0.038;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9  
| | | | |  
Db 30 ILDDIGHGV 38

```
RESULT 2
US-09-051-982A-13
; Sequence 13, Application US/09051982A
; Patent No. 6255074
; GENERAL INFORMATION:
; APPLICANT: PENDERGAST, ANN MARIE
; APPLICANT: DAI, ZONGHAN
; TITLE OF INVENTION: ABL-INTERACTOR PROTEIN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/051.982A
; FILING DATE: 08-JUL-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-193
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; MOLECULE TYPE: protein
; US-09-051-982A-13

Query Match 100.0%; Score 48; DB 3; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9
Db 30 ILDDIGHGV 38

RESULT 3
US-08-549-004A-5
; Sequence 5, Application US/08549004A
; Patent No. 5969101
; GENERAL INFORMATION:
; APPLICANT: PENDERGAST, ANN MARIE
; APPLICANT: DAI, ZONGHAN
; TITLE OF INVENTION: ABL-INTERACTOR PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/549,004A

Query Match 100.0%; Score 48; DB 3; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9
Db 30 ILDDIGHGV 38

RESULT 4
US-09-051-982A-5
; Sequence 5, Application US/09051982A
; Patent No. 6255074
; GENERAL INFORMATION:
; APPLICANT: PENDERGAST, ANN MARIE
; APPLICANT: DAI, ZONGHAN
; TITLE OF INVENTION: ABL-INTERACTOR PROTEIN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/051.982A
; FILING DATE: 08-JUL-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-193
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: linear
; US-09-051-982A-5

Query Match 100.0%; Score 48; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9
Db 100 ILDDIGHGV 108
```



Qy 1 ILDDIGHGV 9  
Db 100 ILDDIGHGV 108

RESULT 5  
US-08-665-647-11  
; Sequence 11, Application US/08665647  
; Patent No. 5935803  
; GENERAL INFORMATION:  
; APPLICANT: Dasquez, Nicki J.  
; APPLICANT: Ron, Dorit  
; APPLICANT: Voronova, Anna F.  
; APPLICANT: Napolitano, Eugene W.  
; TITLE OF INVENTION: METHODS TO IDENTIFY IMMUNOMODULATORS  
; USING COGNATE INTERACTION OF PKC-THETA  
; NUMBER OF SEQUENCES: 89  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Avenue, NW - Ste. 5500  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1888  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/665,647  
; APPLICATION NUMBER: US/08/665,647  
; FILING DATE: 18-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murashige, Kate H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 22550-20025.25  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 822-0168  
; TELEX: 90-4030 MRSNFOERSWSH  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-665-647-11

Query Match 95.8%; Score 46; DB 2; Length 239;  
Best Local Similarity 77.8%; Pred. No. 0.54;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9  
Db 101 VLDDVGHGV 109

RESULT 6  
US-09-189-035-5  
; Sequence 5, Application US/09189035  
; Patent No. 6020165  
; GENERAL INFORMATION:  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Baughn, Mariah R.  
; TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS  
; FILE REFERENCE: PF-0638 US  
; CURRENT APPLICATION NUMBER: US/09/189,035  
; CURRENT FILING DATE: 1998-11-10  
; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PERL Program  
; SEQ ID NO 5  
; LENGTH: 480  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: g2245671  
US-09-189-035-5

Query Match 95.8%; Score 46; DB 3; Length 480;  
Best Local Similarity 77.8%; Pred. No. 1.2;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9  
Db 145 VLDDVGHGV 153

RESULT 7  
US-09-382-086-5  
; Sequence 5, Application US/09382086  
; Patent No. 6201106  
; GENERAL INFORMATION:  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Baughn, Mariah R.  
; TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS  
; FILE REFERENCE: PF-0638 US  
; CURRENT APPLICATION NUMBER: US/09/382,086  
; CURRENT FILING DATE: 1998-08-24  
; EARLIER APPLICATION NUMBER: 09/189,035  
; EARLIER FILING DATE: 1998-11-10  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PERL Program  
; SEQ ID NO 5  
; LENGTH: 480  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: g2245671  
US-09-382-086-5

Query Match 95.8%; Score 46; DB 3; Length 480;  
Best Local Similarity 77.8%; Pred. No. 1.2;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9  
Db 145 VLDDVGHGV 153

RESULT 8  
US-09-949-016-8074  
; Sequence 8074, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8074  
; LENGTH: 531

```

; TYPE: PRT
; ORGANISM: Human
; US-09-949-016--8074

Query Match          95.8%; Score 46; DB 4; Length 531;
Best Local Similarity 77.8%; Pred. No. 1.3;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9
Db 170 VLDDVGHGV 178
      :|||:||||

RESULT 9
US-09-377-285B-20
; Sequence 20, Application US/09377285B
; Patent No. 6720175
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: WORLEY, Paul
; APPLICANT: TU, Jian
; APPLICANT: XIAO, Bo
; APPLICANT: LEAHY, Daniel
; APPLICANT: BENEKEN, Jutta
; APPLICANT: LANAHAN, Anthony
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)
; FILE REFERENCE: JHU1580-4
; CURRENT APPLICATION NUMBER: US/09/377,285B
; CURRENT FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: US 60/138,426
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,493
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,494
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/097,334
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-377-285B-20

Query Match          89.6%; Score 43; DB 4; Length 366;
Best Local Similarity 87.5%; Pred. No. 3.2;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDDIGHGV 9
Db 146 LDDIGHGI 153
      :|||:||||:

RESULT 10
US-09-377-285B-16
; Sequence 16, Application US/09377285B
; Patent No. 6720175
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: WORLEY, Paul
; APPLICANT: TU, Jian
; APPLICANT: XIAO, Bo
; APPLICANT: LEAHY, Daniel
; APPLICANT: BENEKEN, Jutta
; APPLICANT: LANAHAN, Anthony
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)
; FILE REFERENCE: JHU1580-4
; CURRENT APPLICATION NUMBER: US/09/377,285B
; CURRENT FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: US 60/138,426
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,493

```

;  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: 2280326  
US-09-382-086-1

Query Match 89.6%; Score 43; DB 3; Length 476;  
Best Local Similarity 87.5%; Pred. No. 4.2;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDDIGHGV 9  
Db 256 LDDIGHGI 263

## RESULT 13

US-09-252-991A-27032  
; Sequence 27032, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27032  
; LENGTH: 552  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27032

Query Match 83.3%; Score 40; DB 4; Length 552;

Best Local Similarity 75.0%; Pred. No. 18;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILDDIGHG 8  
Db 426 LDDIGHG 433

## RESULT 14

US-09-252-991A-20079  
; Sequence 20079, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20079  
; LENGTH: 362  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (37), (126), (151)  
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.  
US-09-252-991A-20079

Query Match 79.2%; Score 38; DB 4; Length 362;

Best Local Similarity 62.5%; Pred. No. 27;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILDDIGHG 8  
Db 283 VLDEVGHG 290

## RESULT 15

US-09-252-991A-23024  
; Sequence 23024, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23024  
; LENGTH: 483  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23024

Query Match 79.2%; Score 38; DB 4; Length 483;

Best Local Similarity 62.5%; Pred. No. 37;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILDDIGHG 8  
Db 404 VLDEVGHG 411

Search completed: November 18, 2005, 09:35:59  
Job time : 14.6452 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 01:08:37 ; Search time 44.3226 Seconds  
(without alignments)  
84.961 Million cell updates/sec

Title: US-10-006-177-7

Perfect score: 48

Sequence: 1 ILDDIGHGV 9

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	9	14	US-10-006-177-7
2	48	100.0	475	18	US-10-756-149-5477
3	46	95.8	173	14	US-10-106-698-4392
4	43	89.6	366	14	US-10-192-381-20
5	43	89.6	366	20	US-11-008-889-20
6	43	89.6	370	14	US-10-192-381-16
7	43	89.6	370	20	US-11-008-889-16
8	40	83.3	473	20	US-11-097-143-20694
9	39	81.2	259	15	US-10-425-114-51295
10	39	81.2	380	18	US-10-450-763-37847
11	39	81.2	500	15	US-10-425-114-70502

12	39	81.2	629	18	US-10-450-763-34877	Sequence 34877, A
13	37	77.1	154	16	US-10-437-963-196639	Sequence 196639, A
14	36	75.0	151	16	US-10-425-115-342550	Sequence 342550, A
15	36	75.0	185	16	US-10-767-701-57517	Sequence 57517, A
16	36	75.0	378	15	US-10-425-114-39429	Sequence 39429, A
17	36	75.0	402	16	US-10-425-115-356942	Sequence 356942, A
18	36	75.0	416	17	US-10-469-204-144	Sequence 144, App
19	36	75.0	746	16	US-10-437-963-146802	Sequence 146802, A
20	35	72.9	49	16	US-10-437-963-106596	Sequence 106596, A
21	35	72.9	68	14	US-10-029-386-28599	Sequence 28599, A
22	35	72.9	88	9	US-09-727-739B-5	Sequence 5, Appli
23	35	72.9	114	9	US-09-727-739B-3	Sequence 3, Appli
24	35	72.9	133	16	US-10-767-701-38338	Sequence 38338, A
25	35	72.9	160	15	US-10-424-599-217079	Sequence 217079, A
26	35	72.9	174	15	US-10-425-114-52263	Sequence 52263, A
27	35	72.9	219	16	US-10-425-115-212629	Sequence 212629, A
28	35	72.9	262	20	US-11-011-913-1	Sequence 1, Appli
29	35	72.9	273	16	US-10-767-701-41564	Sequence 41564, A
30	35	72.9	409	15	US-10-424-599-238679	Sequence 238679, A
31	35	72.9	420	20	US-11-097-143-42174	Sequence 42174, A
32	35	72.9	449	9	US-09-738-626-5315	Sequence 5315, Ap
33	35	72.9	449	18	US-10-501-282-136	Sequence 136, App
34	35	72.9	473	16	US-10-437-963-186053	Sequence 186053, A
35	35	72.9	473	16	US-10-739-930-6388	Sequence 6388, Ap
36	35	72.9	484	17	US-10-494-675-78	Sequence 78, Appl
37	35	72.9	488	16	US-10-425-115-324281	Sequence 324281, A
38	35	72.9	518	15	US-10-425-114-59128	Sequence 59128, A
39	35	72.9	526	18	US-10-450-763-45798	Sequence 45798, A
40	35	72.9	585	15	US-10-425-114-53676	Sequence 53676, A
41	35	72.9	693	15	US-10-425-114-56454	Sequence 56454, A
42	35	72.9	770	16	US-10-437-963-166183	Sequence 166183, A
43	35	72.9	788	16	US-10-437-963-145665	Sequence 145665, A
44	35	72.9	815	16	US-10-425-115-231132	Sequence 231132, A
45	35	72.9	906	15	US-10-080-334-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1  
US-10-006-177-7  
; Sequence 7, Application US/10006177  
; Publication No. US20030165513A1  
; GENERAL INFORMATION:  
; APPLICANT: Ramakrishna, Venky  
; APPLICANT: Ross, Mark  
; APPLICANT: Philip, Ramila  
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treatm  
; FILE REFERENCE: 26747-35  
; CURRENT APPLICATION NUMBER: US/10/006,177  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: US/60/251,022  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: US/60/256,824  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Epitopic Peptide  
US-10-006-177-7

Query Match 100.0%; Score 48; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILDDIGHGV 9

Db 1 ILDDIGHGV 9

```
RESULT 2
US-10-756-149-5477
; Sequence 5477, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10756.149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5477
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-5477

Query Match      100.0%; Score 48; DB 18; Length 475;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ILDDIGHGV 9
Db      145 ILDDIGHGV 153

RESULT 3
US-10-106-698-4392
; Sequence 4392, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4392
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (146)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (155)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (163)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (172)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4392
```

```
Query Match      95.8%; Score 46; DB 14; Length 173;
Best Local Similarity 77.8%; Pred. No. 1.8;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ILDDIGHGV 9
Db      29 VLDDVGHGV 37

RESULT 4
US-10-192-381-20
; Sequence 20, Application US/10192381
; Publication No. US20030170807A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: WORLEY, Paul
; APPLICANT: TU, Jian
; APPLICANT: XIAO, Bo
; APPLICANT: LEAHY, Daniel
; APPLICANT: BENEKEN, Jutta
; APPLICANT: LANAHAN, Anthony
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS
; FILE REFERENCE: JHU1580-4
; CURRENT APPLICATION NUMBER: US/10/192,381
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/377,285
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: US 60/138,426
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,493
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,494
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/097,334
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-192-381-20

Query Match      89.6%; Score 43; DB 14; Length 366;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 LDDIGHGV 9
Db      146 LDDIGHGI 153

RESULT 5
US-11-008-889-20
; Sequence 20, Application US/11008889
; Publication No. US20050164344A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: WORLEY, Paul
; APPLICANT: TU, Jian
; APPLICANT: XIAO, Bo
; APPLICANT: LEAHY, Daniel
; APPLICANT: BENEKEN, Jutta
; APPLICANT: LANAHAN, Anthony
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS
; FILE REFERENCE: JHU1580-4
; CURRENT APPLICATION NUMBER: US/11/008,889
; CURRENT FILING DATE: 2004-12-10
; PRIOR APPLICATION NUMBER: US/10/192,381
; PRIOR FILING DATE: 2002-07-09
```

; PRIOR APPLICATION NUMBER: US/09/377,285  
; PRIOR FILING DATE: 1999-08-18  
; PRIOR APPLICATION NUMBER: US 60/138,426  
; PRIOR FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: US 60/138,493  
; PRIOR FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: US 60/138,494  
; PRIOR FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: US 60/097,334  
; PRIOR FILING DATE: 1998-08-18  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 20  
; LENGTH: 366  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-008-889-20

Query Match 89.6%; Score 43; DB 20; Length 366;  
Best Local Similarity 87.5%; Pred. No. 14;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDDIGHGV 9  
|||:||||  
Db 146 LDDIGHGI 153

RESULT 6  
US-10-192-381-16  
; Sequence 16, Application US/10192381  
; Publication No. US20030170807A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: WORLEY, Paul  
; APPLICANT: XIAO, Jian  
; APPLICANT: LEAHY, Daniel  
; APPLICANT: BENEKEN, Jutta  
; APPLICANT: LANAHAN, Anthony  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS FILED)  
; FILE REFERENCE: JHU1580-4  
; CURRENT APPLICATION NUMBER: US/10/192,381  
; CURRENT FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: US/09/377,285  
; PRIOR FILING DATE: 1999-08-18  
; PRIOR APPLICATION NUMBER: US 60/138,426  
; PRIOR FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: US 60/138,493  
; PRIOR FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: US 60/138,494  
; PRIOR FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: US 60/097,334  
; PRIOR FILING DATE: 1998-08-18  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 370  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-192-381-16

Query Match 89.6%; Score 43; DB 14; Length 370;  
Best Local Similarity 87.5%; Pred. No. 14;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDDIGHGV 9  
|||:||||  
Db 152 LDDVGHGV 159

RESULT 7  
US-11-008-889-16

; Sequence 16, Application US/11008889  
; Publication No. US20050164344A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: WORLEY, Paul  
; APPLICANT: TU, Jian  
; APPLICANT: XIAO, Bo  
; APPLICANT: LEAHY, Daniel  
; APPLICANT: BENEKEN, Jutta  
; APPLICANT: LANAHAN, Anthony  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS FILED)  
; FILE REFERENCE: JHU1580-4  
; CURRENT APPLICATION NUMBER: US/11/008,889  
; CURRENT FILING DATE: 2004-12-10  
; PRIOR APPLICATION NUMBER: US/10/192,381  
; PRIOR FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: US/09/377,285  
; PRIOR FILING DATE: 1999-08-18  
; PRIOR APPLICATION NUMBER: US 60/138,426  
; PRIOR FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: US 60/138,493  
; PRIOR FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: US 60/138,494  
; PRIOR FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: US 60/097,334  
; PRIOR FILING DATE: 1998-08-18  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 370  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-11-008-889-16

Query Match 89.6%; Score 43; DB 20; Length 370;  
Best Local Similarity 87.5%; Pred. No. 14;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDDIGHGV 9  
|||:||||  
Db 152 LDDVGHGV 159

RESULT 8  
US-11-097-143-20694  
; Sequence 20694, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23

; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20694  
; LENGTH: 473  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-20694

Query Match 83.3%; Score 40; DB 20; Length 473;  
Best Local Similarity 66.7%; Pred. No. 67;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILDDIGHGV 9  
:|||||:  
Db 157 MLDEIGHGI 165

RESULT 9  
US-10-425-114-51295  
; Sequence 51295, Application US/10425114  
; Publication No. US20040034889A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 51295  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700071826\_FLI.pep  
US-10-425-114-51295

Query Match 81.2%; Score 39; DB 15; Length 259;  
Best Local Similarity 85.7%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DDIGHGV 9  
:|||||:  
Db 249 DDVGHGV 255

RESULT 10  
US-10-450-763-37847  
; Sequence 37847, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 37847  
; LENGTH: 380  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-450-763-37847

Query Match 81.2%; Score 39; DB 18; Length 380;  
Best Local Similarity 66.7%; Pred. No. 81;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILDDIGHGV 9  
:|||||:  
Db 267 ILHDVGHGI 275

RESULT 11

US-10-425-114-70502  
; Sequence 70502, Application US/10425114  
; Publication No. US20040034889A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 70502  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-2MFLM017036H04\_FLI.pep  
US-10-425-114-70502

Query Match 81.2%; Score 39; DB 15; Length 500;  
Best Local Similarity 85.7%; Pred. No. 11e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DDIGHGV 9  
:|||||:  
Db 249 DDVGHGV 255

RESULT 12

US-10-450-763-34877  
; Sequence 34877, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 34877  
; LENGTH: 629  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (512)..(544)  
; OTHER INFORMATION: ABC transporters family proteins domain identified by  
; OTHER INFORMATION: eMATRIX, accession number BL00211B, p-value=3.423e-17, raw score  
; OTHER INFORMATION: 13.37



```

; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (436)...(588)
; OTHER INFORMATION: ABC transporter domain identified by PFam, accession name
; OTHER INFORMATION: ABC_tran, E-value=2e-07, PFam score of 26.5
US-10-450-763-34877

Query Match      81.2%; Score 39; DB 18; Length 629;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9
Db 226 ILHDVGHGI 234

RESULT 13
US-10-437-963-196639
; Sequence 196639, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 196639
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_92471C.1.pgp
US-10-437-963-196639

Query Match      77.1%; Score 37; DB 16; Length 154;
Best Local Similarity 75.0%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 8
Db 34 ILDDAGHG 41

RESULT 14
US-10-425-115-342550
; Sequence 342550, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 342550
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_75572C.1.pgp

```

```

US-10-425-115-342550

Query Match      75.0%; Score 36; DB 16; Length 151;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 8
Db 20 VLQDVGHG 27

RESULT 15
US-10-767-701-57517
; Sequence 57517, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 57517
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30968284.pgp
US-10-767-701-57517

Query Match      75.0%; Score 36; DB 16; Length 185;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILDDIGHGV 9
Db 74 VVTDVGHGV 82

Search completed: November 18, 2005, 10:44:39
Job time : 46.3226 secs

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:03:06 ; Search time 48.7742 Seconds  
(without alignments)  
71.366 Million cell updates/sec

Title: US-10-006-177-8

Perfect score: 42

Sequence: 1 LLDRFLATV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	5	ABG32308 HLA-A2 as
2	42	100.0	9	6	ABJ19986 MHC bindi
3	42	100.0	10	6	ABJ20156 MHC bindi
4	42	100.0	300	8	ADN01131 Human cel
5	42	100.0	377	2	Aaw21965 Human cyc
6	42	100.0	377	3	Aaw52185 Human cyc
7	42	100.0	377	6	ABR3934 Human pro
8	42	100.0	377	8	ADP12615 Protein e
9	42	100.0	377	8	ABM80569 Tumour-as
10	35	83.3	303	7	ADC95958 E. faeciu
11	34	81.0	561	7	ADC95696 E. faeciu
12	34	81.0	3312	6	ABM66964 Photorhab
13	34	81.0	5217	6	ABM70675 Photorhab
14	34	81.0	5458	6	ABM66970 Photorhab
15	34	81.0	16368	6	ABM67171 Photorhab
16	33	78.6	239	4	AAM16821 Peptide #
17	33	78.6	239	4	ABB35805 Peptide #
18	33	78.6	239	4	AAM29304 Peptide #
19	33	78.6	239	4	ABB30636 Peptide #
20	33	78.6	239	4	ABB23116 Protein #
21	33	78.6	239	4	AAM70981 Human bon
22	33	78.6	239	4	AAM56613 Human liv
23	33	78.6	239	4	ABG50656 Human pep
24	33	78.6	239	5	ABG38581 Thermococ
25	33	78.6	273	8	ADN46132

## ALIGNMENTS

### RESULT 1

ABG32308  
ID ABG32308 standard; peptide; 9 AA.

XX AC ABG32308;

XX DT 05-NOV-2002 (first entry)

XX DE HLA-A2 associated immunogenic peptide from human Cyclin protein.

XX KW Human; immunogen; epitope; HLA-A1; human leukocyte antigen; CTL;  
XX KW cytotoxic Y lymphocyte; cytostatic; cancer; colorectal carcinoma;  
XX KW ovarian carcinoma; lung carcinoma; prostate carcinoma; vaccine; tumour;  
XX KW HLA-2; passive immunotherapy; cyclin.

XX OS Homo sapiens.

XX PN WO200246416-A2.

XX PD 13-JUN-2002.

XX PF 04-DEC-2001; 2001WO-US047290.

XX PR 04-DEC-2000; 2000US-0251022P.

XX PA 20-DEC-2000; 2000US-0256824P.

XX PI (ARGO-) ARGONEX INC.

XX PI Ramakrishna V, Ross M, Philip R;

XX DR WPI; 2002-619021/66.

XX PT New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte,  
XX PT and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.

XX PS Claim 1; Page 50; 60pp; English.

XX CC The invention relates to an immunogen comprising an isolated polypeptide  
XX CC whose amino acid sequence comprises an epitopic peptide, does not include  
XX CC MAGE 4 or MFG-E8 proteins, or consists of MAGE D protein or its  
XX CC immunologically active fragment. Also included are a polynucleotide  
XX CC encoding the immunogen or its complement, a vector comprising the  
XX CC polynucleotide, a mammalian cell comprising the vector and expressing the  
XX CC antibody specific for the immunogen. The immunogen is useful for inducing  
XX CC a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour  
XX CC cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is  
XX CC useful for inducing a CTL response when administered to a subject. A

Abm82315 Tumour-as  
Adc51716 Crab-eati  
Abw01475 Monkey MC  
Adh53374 Cynomolgu  
Abw01478 Monkey MC  
Abw01473 Monkey MC  
Adh53370 Cynomolgu  
Aaw61370 Human 11C  
Aay16570 Amino aci  
Aab13442 Truncated  
Aab13437 Rat MCH1  
Aay57284 Human GPC  
Aab12778 Rat SLC-1  
Aay97310 Murine 11  
Aay90258 Human 11C  
Aae04747 Human 11C  
Aay97668 Human MCH  
Aae07328 Human mel  
Aab96870 Rat SLC-1  
Aab85895 Human sho

26 33 78.6 303 8 ABM82315  
27 33 78.6 337 7 ADC51716  
28 33 78.6 345 7 ABW01475  
29 33 78.6 345 7 ADH53374  
30 33 78.6 347 7 ABW01478  
31 33 78.6 352 7 ABW01473  
32 33 78.6 352 7 ADH53370  
33 33 78.6 353 2 AAW61370  
34 33 78.6 353 3 AAY16570  
35 33 78.6 353 3 AAB13442  
36 33 78.6 353 3 AAB13437  
37 33 78.6 353 3 AAY57284  
38 33 78.6 353 3 AAB12778  
39 33 78.6 353 3 AAY97310  
40 33 78.6 353 3 AAY90258  
41 33 78.6 353 4 AAE04747  
42 33 78.6 353 4 AAY97668  
43 33 78.6 353 4 AAE07328  
44 33 78.6 353 4 AAB96870  
45 33 78.6 353 4 AAB85895



SQ Sequence 10 AA;  
 Query Match 100.0%; Score 42; DB 6; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.086;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLDRFLATV 9  
 DB 2 LLDRFLATV 10  
 |||||

RESULT 4  
 ID ADN01131 standard; protein; 300 AA.  
 XX  
 AC ADN01131;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Human cell growth, differentiation, and death-associated protein #15.  
 KW human; cell growth; cell differentiation; cell death; CGDD;  
 KW cell proliferative disorder; arteriosclerosis; atherosclerosis; bursitis;  
 KW cirrhosis; hepatitis; polycythaemia vera; psoriasis; cancer;  
 KW developmental disorder; Cushing's syndrome; hypothyroidism;  
 KW neurological disorder; epilepsy; stroke; Alzheimer's disease;  
 KW Pick's disease; Huntington's disease; Parkinson's disease;  
 KW multiple sclerosis; autoimmune disorder; inflammatory disorder; AIDS;  
 KW allergy; anaemia; asthma; contact dermatitis; diabetes mellitus;  
 KW reproductive disorder; infertility; endometriosis; uterine fibroid.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2004031364-A2.  
 PN  
 XX  
 PD 15-APR-2004.  
 XX  
 PF 03-OCT-2003; 2003WO-US031441.  
 XX  
 PR 03-OCT-2002; 2002US-0416205P.  
 PR 25-OCT-2002; 2002US-0421521P.  
 PR 21-NOV-2002; 2002US-0428376P.  
 PR 23-DEC-2002; 2002US-0436258P.  
 PR 10-JAN-2003; 2003US-0439292P.  
 PR 13-FEB-2003; 2003US-0447578P.  
 XX  
 PA (INCY-) INCYTE CORP.  
 PA (BURR/) BURRILL J D.  
 XX  
 PI Elliott VS, Swarnakar A, Tang YT, Yue H, Lu DAM, Khare R;  
 PI Chawla NK, Richardson TW, Marquis JP, Lal PG, Nguyen DB, Lee SY;  
 PI Tran YG, Bhatia UG, Lee S, Blake JJ, Ho A, Zheng W, Gao J, Tran B;  
 PI Yang YK, Gietzen KJ, Hafalia AJA;  
 XX  
 XX WPI; 2004-330172/30.  
 DR N-PSDB; ADN01156.  
 XX  
 PT New isolated polypeptides associated with cell growth, differentiation  
 PT and death, useful for diagnosing, treating or preventing e.g.  
 PT atherosclerosis, psoriasis, cancers, Alzheimer's disease, AIDS, anemia,  
 PT diabetes mellitus or infertility.  
 XX  
 PS Claim 1; SEQ ID NO 15; 213pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of human  
 CC proteins that are associated with cell growth, differentiation, and death  
 CC (CGDD). The DNA and protein sequences of the invention are useful for  
 CC diagnosing, treating or preventing disorders associated with aberrant  
 CC expression of CGDD, such as: cell proliferative disorders (e.g.  
 CC arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis,  
 CC polycythaemia vera, psoriasis and cancers), developmental disorders (e.g.  
 CC Cushing's syndrome and hypothyroidism), neurological disorders (e.g.  
 CC epilepsy, stroke, Alzheimer's disease, Pick's disease, Huntington's

CC disease, Parkinson's disease and multiple sclerosis),  
 CC autoimmune/inflammatory disorders (e.g. AIDS, allergies, anaemia, asthma,  
 CC contact dermatitis and diabetes mellitus), and reproductive disorders  
 CC (e.g. infertility, endometriosis and uterine fibroid). The present amino  
 CC acid sequence represents a human CGDD-associated protein of the  
 CC invention.  
 XX  
 SQ Sequence 300 AA;  
 Query Match 100.0%; Score 42; DB 8; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLDRFLATV 9  
 DB 72 LLDRFLATV 80  
 |||||

RESULT 5  
 AAW21965  
 ID AAW21965 standard; protein; 377 AA.  
 XX  
 AC AAW21965;  
 XX  
 DT 02-DEC-1997 (first entry)  
 XX  
 DE Human cyclin I.  
 KW human; cyclin I; antisense; probe; neurone; cancer; antibody.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9712973-A1.  
 PN  
 XX  
 PD 10-APR-1997.  
 XX  
 PF 07-OCT-1996; 96WO-JP002905.  
 XX  
 PR 05-OCT-1995; 95JP-00284663.  
 XX  
 PA (SUME) SUMITOMO ELECTRIC IND CO.  
 XX  
 PI Nakamura T;  
 XX  
 XX WPI; 1997-226217/20.  
 DR N-PSDB; AAT73937.  
 XX  
 PT Human cyclin I protein and related (anti:sense) DNA - used for neuron  
 PT labelling method and cancer cell detection.  
 XX  
 PS Claim 1; Fig 1; 45pp; Japanese.  
 XX  
 CC This sequence is human cyclin I. Antisense polynucleotides are useful for  
 CC as probes and can be labelled and used for detection of neurones by  
 CC hybridisation with mRNA for cyclin I (contained in the neurones and  
 CC arising by the expression of the cyclin I gene in these cells). The gene  
 CC can be used for detection of cancer cells by detecting the expression of  
 CC the cyclin I gene in these cells. Also antibodies specific for the  
 CC fragments of the protein (especially AAW21966) can be used for detection  
 XX  
 SQ Sequence 377 AA;  
 Query Match 100.0%; Score 42; DB 2; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 4.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLDRFLATV 9  
 DB 72 LLDRFLATV 80  
 |||||

RESULT 6  
 AAY52185

ID AAY52185 standard; protein; 377 AA.  
XX  
AC AAY52185;  
XX  
DT 09-FEB-2000 (first entry)  
XX  
DE Human cyclin I amino acid sequence.  
XX  
KW Cyclin-dependent kinase 2; CDK2; hsReq; hsReq\*-1; hsReq\*-2; cyclin I;  
KW ERH; cell cycle; proliferation; cancer; hyperproliferative disorder;  
KW atherosclerosis; tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO9925829-A2.  
XX  
PD 27-MAY-1999.  
XX  
PF 12-NOV-1998; 98WO-US024095.  
XX  
PR 13-NOV-1997; 97US-00969106.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Yang M, Nandabalan K, Schulz VP;  
XX  
DR WPI; 2000-061923/05.  
XX  
DR N-PSDB; AAZ37836.  
XX  
PT New complexes of the cyclin-dependent kinase 2 protein with its  
XX interacting proteins, used to treat, e.g. atherosclerosis.  
XX  
PS Example; Fig 2; 90pp; English.  
XX  
CC This is the human Cyclin I amino acid sequence. Cyclin I is expressed at  
CC almost constant levels throughout the cell cycle, and is implicated in  
CC controlling cell cycle progression and transcriptional control. Cyclins  
CC form complexes with cyclin-dependent kinases. CDK2, cyclin-dependent  
CC kinase 2, is used in the invention which relates to complexes of the CDK2  
CC protein with other proteins, selected from cyclin I, ERH, hReq, hReq\*-1  
CC and hReq\*-2 (AAY52185-Y52188). CDK2 is expressed late in G1 or early in  
CC S phase of the cell cycle, and is pivotal for G1/S transition.  
CC Compositions containing a CDK2 complex, an antibody targeting the  
CC complex, and nucleotide sequences encoding CDK2 or its derivatives can be  
CC used therapeutically. The complexes and their nucleotide sequences can be  
CC used to treat diseases or disorders associated with increased or  
CC decreased levels of the complex. Screening the complex, or a derivative  
CC or a modulator of the complex for neoplastic activity by measuring the  
CC survival or proliferation of cells from a malignant cell line when in  
CC contact with the complex can be used to indicate if the complex has  
CC anti-neoplastic activity. Screening for molecules that modulate the  
CC formation of the complexes can be used for treating or preventing  
CC atherosclerosis or atherosclerosis-associated disease by contacting cells  
CC or using a test animal, in which tumour growth or regression is measured  
CC to test whether anti-neoplastic activity is displayed. Diseases which can  
CC be treated or prevented by molecule/s which modulate the function of the  
CC complex include cancer, hyperproliferative disorders and atherosclerosis  
XX  
SQ Sequence 377 AA;  
Query Match 100.0%; Score 42; DB 3; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LLDRFLATV 9  
Db 72 LLDRFLATV 80  
RESULT 7  
ABR39934  
ID ABR39934 standard; protein; 377 AA.  
XX

AC ABR39934;  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Human prostate selective polypeptide Pr325.  
XX  
KW Prostate; molecular marker; cancer; cytostatic; gene therapy; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2003014298-A2.  
XX  
PD 20-FEB-2003.  
XX  
PF 02-AUG-2002; 2002WO-US024431.  
XX  
PR 03-AUG-2001; 2001US-0309470P.  
XX  
PR 30-OCT-2001; 2001US-0330747P.  
XX  
PA (ORIG-) ORIGENE TECHNOLOGIES INC.  
XX  
PI Sun Z, Li X, Jay G, Kovacs KP, Shu Y, Fan W;  
XX  
DR WPI; 2003-256562/25.  
XX  
DR N-PSDB; ACC47339.  
XX  
PT New polynucleotide, useful for preparing a composition for treating  
XX prostate disease, e.g., cancer.  
XX  
PS Claim 5; Page 147-149; 212pp; English.  
XX  
CC The invention relates to prostate selective polynucleotides and  
XX polypeptides. The polynucleotides are expressed in prostate and are  
XX useful as molecular markers, as drug targets, and for detecting,  
XX monitoring, preventing or treating diseases and conditions related to  
XX prostate, such as prostate cancers. The present sequence represents a  
XX prostate specific polypeptide  
XX  
SQ Sequence 377 AA;  
Query Match 100.0%; Score 42; DB 6; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LLDRFLATV 9  
Db 72 LLDRFLATV 80  
RESULT 8  
ADP12615  
ID ADP12615 standard; protein; 377 AA.  
XX  
AC ADP12615;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Protein encoded by mRNA of the invention #225.  
XX  
KW transplant rejection; immune system; rheumatoid arthritis; lupus;  
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.  
XX  
OS Homo sapiens.  
XX  
PN WO2004042346-A2.  
XX  
PD 21-MAY-2004.  
XX  
PF 24-APR-2003; 2003WO-US012946.  
XX  
PR 24-APR-2002; 2002US-00131831.  
XX  
PR 20-DEC-2002; 2002US-00325899.  
XX

PA (EXPR-) EXPRESSION DIAGNOSTICS INC.  
 XX Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;  
 PI Rosenberg S;  
 XX WPI; 2004-400724/37.  
 DR  
 XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,  
 PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant  
 PT rejection, in an individual, comprises detecting the expression level of  
 PT the genes.  
 XX  
 XX Claim 65; SEQ ID NO 2624; 1762pp; English.  
 PS  
 CC The present invention relates to diagnosing or monitoring transplant  
 CC rejection, e.g. cardiac or kidney transplant rejection, in an individual  
 CC comprising detecting the expression level of one or more genes. The  
 CC methods, system and kits are useful in diagnosing or monitoring  
 CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic  
 CC islet, lung, bone marrow or stem cell transplant rejection,  
 CC xenotransplant rejection or mechanical organ replacement rejection, in an  
 CC individual. The method is also useful in assessing the immune status of  
 CC an individual. The methods are also useful in diagnosing and monitoring  
 CC diseases that involve the immune system, e.g. rheumatoid arthritis,  
 CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or  
 CC viral, bacterial or fungal infection. The present sequence represents a  
 CC protein that is encoded by the mRNA of the invention.  
 XX  
 XX Sequence 377 AA;  
 SQ

Query Match 100.0%; Score 42; DB 8; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 4.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 LLDRFLATV 9  
 Db 72 LLDRFLATV 80

RESULT 9  
 ABM80569  
 ID ABM80569 standard; protein; 377 AA.  
 XX  
 AC ABM80569;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Tumour-associated antigenic target (TAT) polypeptide PRO4605, SEQ.1450.  
 XX  
 KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;  
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
 KW central nervous system cancer; bladder cancer; pancreatic cancer;  
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;  
 KW chromosome identification; chromosome mapping; gene mapping;  
 KW gene therapy; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004030615-A2.  
 XX  
 PD 15-APR-2004.  
 XX  
 PF 29-SEP-2003; 2003WO-US028547.  
 XX  
 PR 02-OCT-2002; 2002US-0414971P.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 XX Wu TD, Zhang Z, Zhou Y;  
 PI WPI; 2004-347921/32.  
 DR N-PSDB; ACN38124.  
 XX  
 PT New isolated nucleic acid derived from Enterococcus faecium encoding an  
 PT Enterococcus faecium polypeptide useful for detection, prevention and

XX New tumor-associated antigenic target polypeptides and nucleic acids,  
 PT useful in preparing a medicament for treating or detecting a  
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
 PT prostate cancer or tumor.  
 XX  
 PS Claim 12; SEQ ID NO 1450; 7273pp; English.  
 XX  
 CC The invention relates to human tumour-associated antigenic target (TAT)  
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
 CC overexpressed in cancer tissues compared to normal tissues, and may thus  
 CC serve as effective targets for the diagnosis and treatment of cancer in  
 CC mammals. The invention also relates to nucleic acid and polypeptide  
 CC sequences at least 80% identical to the TAT nucleic acids and  
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
 CC TAT polypeptide; and methods and compositions for the treatment or  
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
 CC antibodies, antagonists, binding molecules and compositions are useful  
 CC for diagnosing or treating a cell proliferative disorder associated with  
 CC increased TAT expression, particularly cancers such as breast cancer,  
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
 CC used as hybridisation probes, in chromosome and gene mapping, in  
 CC chromosome identification and in gene therapy. The present sequence  
 CC represents a TAT polypeptide of the invention  
 XX  
 XX Sequence 377 AA;  
 SQ

Query Match 100.0%; Score 42; DB 8; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 4.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 LLDRFLATV 9  
 Db 72 LLDRFLATV 80

RESULT 10  
 ADC95958  
 ID ADC95958 standard; protein; 303 AA.  
 XX  
 AC ADC95958;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE E. faecium protein sequence SEQ ID 5585.  
 XX  
 KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;  
 KW abdominal-pelvic infection.  
 XX  
 OS Enterococcus faecium.  
 XX  
 PN US6583275-B1.  
 XX  
 PD 24-JUN-2003.  
 XX  
 PF 30-JUN-1998; 98US-00107532.  
 XX  
 PR 02-JUL-1997; 97US-0051571P.  
 PR 14-MAY-1998; 98US-0085598P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Doucette-Stamm LA, Bush D;  
 DR WPI; 2003-799836/75.  
 DR N-PSDB; ADC92304.  
 XX  
 PT New isolated nucleic acid derived from Enterococcus faecium encoding an  
 PT Enterococcus faecium polypeptide useful for detection, prevention and

```

PT treatment of a pathological condition resulting from a bacterial
PT infection.
XX
PS Example 1; SEQ ID NO 5585; 243pp; English.
XX
CC The invention relates to an isolated nucleic acid derived from
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridising to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to
CC a single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids are useful for diagnosing pathological conditions
CC resulting from E. faecium bacterial infection (e.g. urinary tract
CC infection), bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of Candida albicans -
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating Enterococcus faecium infections. The present sequence represents
CC one if the disclosed E. faecium proteins.
XX
SQ Sequence 303 AA;

Query Match      83.3%; Score 35; DB 7; Length 303;
Best Local Similarity 77.8%; Pred. No. 87;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLDRFLATV 9
DB 98 LLDRFLVTL 106
      ||||| |
      ||||| |

RESULT 11
ID ADC95696 standard; protein; 561 AA.
XX
AC ADC95696;
XX
XX 01-JAN-2004 (first entry)
XX
DE E. faecium protein sequence SEQ ID 5323.
XX
KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
KW abdominal-pelvic infection.
XX
OS Enterococcus faecium.
XX
PN US6583275-B1.
XX
PD 24-JUN-2003.
XX
PF 30-JUN-1998; 98US-00107532.
XX
PR 02-JUL-1997; 97US-0051571P.
XX
PR 14-MAY-1998; 98US-0085598P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WPI; 2003-799836/75.
DR N-PSDB; ADC92042.
XX
XX New isolated nucleic acid derived from Enterococcus faecium encoding an
PT Enterococcus faecium polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
PT infection.
XX
PS Example 1; SEQ ID NO 5323; 243pp; English.
XX

```

---

```

CC The invention relates to an isolated nucleic acid derived from
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridising to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to
CC a single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids are useful for diagnosing pathological conditions
CC resulting from E. faecium bacterial infection (e.g. urinary tract
CC infection), bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of Candida albicans -
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating Enterococcus faecium infections. The present sequence represents
CC one if the disclosed E. faecium proteins.
XX
SQ Sequence 561 AA;

Query Match      81.0%; Score 34; DB 7; Length 561;
Best Local Similarity 75.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDRFLATV 9
DB 16 IDRFLATV 23
      :|||:|
      :|||:|

RESULT 12
ID AEM66964 standard; protein; 3312 AA.
XX
AC AEM66964;
XX
XX 20-NOV-2003 (first entry)
XX
DE Photorhabdus luminescens protein sequence #61.
XX
KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough.
XX
OS Photorhabdus luminescens.
XX
PN W0200294867-A2.
XX
PD 28-NOV-2002.
XX
PF 07-FEB-2002; 2002WO-IB003040.
XX
PR 07-FEB-2001; 2001FR-00001659.
XX
PA (INSP ) INST PASTEUR.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI Buchrieser C;
XX
XX WPI; 2003-148459/14.
XX
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.
PS Claim 2; SEQ ID NO 61; 1205pp; French.
XX
XX The invention relates to the isolation of genes and their encoded
XX proteins from Photorhabdus luminescens. The isolated sequences are
XX sources of probes and primers for detecting the genome of P. luminescens
XX and related species; to study polymorphisms; for gene analysis and for
XX detection/amplification of the genes. Antibodies (Ab) raised against the

```



CC polypeptides encoded by the genes are used for detection/identification  
 CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that  
 CC carry a gene-containing vector are used to select compounds that  
 CC modulate, regulate, induce or inhibit expression of the genes in plants,  
 CC animals or microorganisms other than *P. luminescens* and are able to alter  
 CC response or sensitivity to toxins and antibiotics produced by *P.*  
 CC *luminescens*. Cells transformed to express the genes are useful for  
 CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically (to treat microbial infection by bacteria or fungi that  
 CC are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which *P.*  
 CC *luminescens* is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated *P. luminescens* proteins  
 XX  
 SQ Sequence 3312 AA;

Query Match 81.0%; Score 34; DB 6; Length 3312;  
 Best Local Similarity 77.8%; Pred. No. 1.8e+03;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLDRLATV 9  
 ||||:|||  
 Db 948 LLDRLAAV 956

RESULT 13  
 ABM70675  
 ID ABM70675 standard; protein; 5217 AA.

XX AC ABM70675;

XX DT 20-NOV-2003 (first entry)

XX DE Photorhabdus luminescens protein sequence #3772.

XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
 XX detection; food; gene expression; plant; animal; microorganism; toxin;  
 XX antibiotic; biopesticide; virulence factor; disease model; plague;  
 XX whooping cough.

XX OS Photorhabdus luminescens.

XX FN WO200294867-A2.

XX PD 28-NOV-2002.

XX PF 07-FEB-2002; 2002WO-IB003040.

XX PR 07-FEB-2001; 2001FR-00001659.

XX PA (INSP ) INST PASTEUR.

XX PA (CNRS ) CNRS CENT NAT RECH SCI.

XX PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
 PI Buchrieser C;

XX DR WPI; 2003-148459/14.

XX PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
 XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX PS Claim 2; SEQ ID NO 3772; 1205pp; French.

XX CC The invention relates to the isolation of genes and their encoded  
 CC proteins from Photorhabdus luminescens. The isolated sequences are  
 CC sources of probes and primers for detecting the genome of *P. luminescens*  
 CC and related species; to study polymorphisms; for gene analysis and for  
 CC detection/amplification of the genes. Antibodies (Ab) raised against the  
 CC polypeptides encoded by the genes are used for detection/identification  
 CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that

CC carry a gene-containing vector are used to select compounds that  
 CC modulate, regulate, induce or inhibit expression of the genes in plants,  
 CC animals or microorganisms other than *P. luminescens* and are able to alter  
 CC response or sensitivity to toxins and antibiotics produced by *P.*  
 CC *luminescens*. Cells transformed to express the genes are useful for  
 CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically (to treat microbial infection by bacteria or fungi that  
 CC are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which *P.*  
 CC *luminescens* is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated *P. luminescens* proteins  
 XX  
 SQ Sequence 5217 AA;

Query Match 81.0%; Score 34; DB 6; Length 5217;  
 Best Local Similarity 77.8%; Pred. No. 2.9e+03;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLDRLATV 9  
 ||||:|||  
 Db 3941 LLDRLAAV 3949

RESULT 14  
 ABM66970  
 ID ABM66970 standard; protein; 5458 AA.

XX AC ABM66970;

XX DT 20-NOV-2003 (first entry)

XX DE Photorhabdus luminescens protein sequence #67.

XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
 XX detection; food; gene expression; plant; animal; microorganism; toxin;  
 XX antibiotic; biopesticide; virulence factor; disease model; plague;  
 XX whooping cough.

XX OS Photorhabdus luminescens.

XX FN WO200294867-A2.

XX PD 28-NOV-2002.

XX PF 07-FEB-2002; 2002WO-IB003040.

XX PR 07-FEB-2001; 2001FR-00001659.

XX PA (INSP ) INST PASTEUR.

XX PA (CNRS ) CNRS CENT NAT RECH SCI.

XX PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
 PI Buchrieser C;

XX DR WPI; 2003-148459/14.

XX PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
 XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX PS Claim 2; SEQ ID NO 67; 1205pp; French.

XX CC The invention relates to the isolation of genes and their encoded  
 CC proteins from Photorhabdus luminescens. The isolated sequences are  
 CC sources of probes and primers for detecting the genome of *P. luminescens*  
 CC and related species; to study polymorphisms; for gene analysis and for  
 CC detection/amplification of the genes. Antibodies (Ab) raised against the  
 CC polypeptides encoded by the genes are used for detection/identification  
 CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that  
 CC carry a gene-containing vector are used to select compounds that  
 CC modulate, regulate, induce or inhibit expression of the genes in plants,

CC animals or microorganisms other than *P. luminescens* and are able to alter  
 CC response or sensitivity to toxins and antibiotics produced by *P.*  
 CC *luminescens*. Cells transformed to express the genes are useful for  
 CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically (to treat microbial infection by bacteria or fungi that  
 CC are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which *P.*  
 CC *luminescens* is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated *P. luminescens* proteins  
 XX  
 SQ Sequence 5458 AA;

Query Match 81.0%; Score 34; DB 6; Length 5458;  
 Best Local Similarity 77.8%; Pred. No. 3e+03;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9  
 ||||:||||  
 Db 952 LLDRYLAHV 960

## RESULT 15

ABM67171  
 ID ABM67171 standard; protein; 16368 AA.

XX AC ABM67171;

XX DT 20-NOV-2003 (first entry)

XX DE Photorhabdus luminescens protein sequence #268.

XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
 XX detection; food; gene expression; plant; animal; microorganism; toxin;  
 XX antibiotic; biopesticide; virulence factor; disease model; plague;  
 XX whooping cough.

XX OS Photorhabdus luminescens.

XX PN WO200294867-A2.

XX PD 28-NOV-2002.

XX PF 07-FEB-2002; 2002WO-IB003040.

XX PR 07-FEB-2001; 2001FR-00001659.

XX PA (INSP ) INST PASTEUR.

XX PA (CNRS ) CNRS CENT NAT RECH SCI.

XX PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
 XX Buchrieser C;

XX PS WPI; 2003-148459/14.

XX PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
 XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX PS Claim 2; SEQ ID NO 268; 1205pp; French.

XX CC The invention relates to the isolation of genes and their encoded  
 CC proteins from Photorhabdus luminescens. The isolated sequences are  
 CC sources of probes and primers for detecting the genome of *P. luminescens*  
 CC and related species; to study polymorphisms; for gene analysis and for  
 CC detection/amplification of the genes. Antibodies (Ab) raised against the  
 CC polypeptides encoded by the genes are used for detection/identification  
 CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that  
 CC carry a gene-containing vector are used to select compounds that  
 CC modulate, regulate, induce or inhibit expression of the genes in plants,  
 CC animals or microorganisms other than *P. luminescens* and are able to alter  
 CC response or sensitivity to toxins and antibiotics produced by *P.*

CC *luminescens*. Cells transformed to express the genes are useful for  
 CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically (to treat microbial infection by bacteria or fungi that  
 CC are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which *P.*  
 CC *luminescens* is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated *P. luminescens* proteins  
 XX  
 SQ Sequence 16368 AA;

Query Match 81.0%; Score 34; DB 6; Length 16368;  
 Best Local Similarity 77.8%; Pred. No. 9.7e+03;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9  
 ||||:||||  
 Db 3388 LLDRYLAHV 3396

Search completed: November 18, 2005, 01:01:20  
 Job time : 53.7742 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:11:06 ; Search time 8.90323 Seconds  
(without alignments)  
97.263 Million cell updates/sec

Title: US-10-006-177-8

Perfect score: 42

Sequence: 1 LLDRFLATV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.79.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	88.1	881	2 C87471	TonB-dependent rec
2	35	83.3	141	1 HAHH	hemoglobin alpha c
3	35	83.3	141	1 HAOEE	hemoglobin alpha c
4	35	83.3	141	1 HARTNG	hemoglobin alpha c
5	35	83.3	141	2 C25929	hemoglobin alpha-I
6	35	83.3	142	1 HAXL1	hemoglobin alpha c
7	35	83.3	142	1 HAXL2	hemoglobin alpha c
8	35	83.3	178	2 JE0264	cyclin I - human
9	34	81.0	197	2 S73688	hypothetical prote
10	33	78.6	116	2 AG3446	arsenate reductase
11	33	78.6	291	2 JC4579	cyclin D2 - chicke
12	33	78.6	320	2 B75496	moXR protein - Dei
13	33	78.6	357	2 A81339	gcpE protein homol
14	33	78.6	422	2 JC7080	melanin-concentrat
15	33	78.6	423	2 JC7677	allatostatin recep
16	33	78.6	449	2 T39410	hypothetical prote
17	33	78.6	599	2 C86161	hypothetical prote
18	33	78.6	676	2 T01084	hypothetical prote
19	33	78.6	950	2 D81821	glycine dehydrogen
20	33	78.6	1436	2 S67655	probable membrane
21	33	78.6	1766	2 E86201	protein F12K11.4 [
22	32	76.2	141	1 HAHOK	hemoglobin alpha c
23	32	76.2	141	1 HAHOC	hemoglobin alpha c
24	32	76.2	141	1 HAHOC	hemoglobin alpha c
25	32	76.2	142	2 S02024	hemoglobin alpha c
26	32	76.2	291	2 S57925	cyclin D2 - Africa
27	32	76.2	298	1 A69879	conserved hypothet
28	32	76.2	352	2 AC2760	sodium bile acid s
29	32	76.2	352	2 A97541	hypothetical prote

30	32	76.2	508	2 S44950	lmbC protein - Str
31	32	76.2	593	1 XYZZFG	frzG protein - Myx
32	32	76.2	617	2 T52499	ferredoxin oxidore
33	32	76.2	822	2 S19947	fibroblast growth
34	32	76.2	822	2 B49151	fibroblast growth
35	31	73.8	133	2 AB3465	hypothetical prote
36	31	73.8	134	2 B90284	hemoglobin alpha-2
37	31	73.8	141	1 HABAG	hemoglobin alpha c
38	31	73.8	141	1 HABAIM	hemoglobin alpha c
39	31	73.8	141	1 HABAY	hemoglobin alpha c
40	31	73.8	141	1 HABTF	hemoglobin alpha c
41	31	73.8	141	1 HABTV	hemoglobin alpha c
42	31	73.8	141	1 HACJB	hemoglobin alpha c
43	31	73.8	141	1 HACZP	hemoglobin alpha c
44	31	73.8	141	1 HADD	hemoglobin alpha c
45	31	73.8	141	1 HAFXB	hemoglobin alpha c

#### ALIGNMENTS

##### RESULT 1

C87471

TonB-dependent receptor [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004

C;Accession: C87471

R;Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: C87471

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-881 <STO>

A;Cross-references: UNIPROT:Q9A7D2; GB:AB005673; NID:gl3423221; PIDN:AAK23767.1; GSPDB:

C;Genetics:

A;Gene: CCI791

Query Match 88.1%; Score 37; DB 2; Length 881;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qv 1 LLDRFLATV 9

:|||:||||

Db 510 MLDRFLATV 518

##### RESULT 2

HAHH

hemoglobin alpha chain - western European hedgehog (tentative sequence)

C;Species: Erinaceus europaeus (western European hedgehog)

C;Date: 31-May-1979 #sequence\_revision 31-May-1979 #text\_change 12-Jul-2004

C;Accession: A02271

R;Maiz, T.; Araya, A.; Matea, G.; Goodman, M.

J. Biochem. 85, 259-269, 1979

A;Title: Amino acid sequences of the alpha and beta chains of adult hemoglobin of the Eu

A;Reference number: A91950; MUID:79109529; PMID:762046

A;Accession: A02271

A;Molecule type: protein

A;Residues: 1-141 <MAI>

A;Cross-references: UNIPROT:P01949

C;Superfamily: globin; globin homology

C;Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carri

F;2-141/Domain: globin homology <GI>

F;58/Binding site: oxygen (His) (distal axial ligand) #status predicted

F;87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 83.3%; Score 35; DB 1; Length 141;  
Best Local Similarity 87.5%; Pred. No. 4;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy      2 LDRFLATV 9
Db      125 LDKFLATV 132

RESULT 3
HAOSE
hemoglobin alpha chain - European mole (tentative sequence)
C/Species: Talpa europaea (European mole)
C/Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 12-Jul-2004
C/Accession: A02273
R/Kleinschmidt, T.; Jellmann, W.; Braunitzer, G.
Hoppe-Seyler's Z. Physiol. Chem. 362, 1263-1272, 1981
A/Title: Die Primärstruktur des Haemoglobins des Maulwurfs (Talpa europaea).
A/Reference number: A91697; MUID:82263326; PMID:7346384
A/Accession: A02273
A/Molecule type: protein
A/Residues: 1-141 <KLE>
A/Cross-references: UNIPROT:P01951
C/Superfamily: globin; globin homology
C/Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrier
F/2-141/Domain: globin homology <GLB>
F/58/Binding site: oxygen (His) (distal axial ligand) #status predicted
F/87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match      83.3%; Score 35; DB 1; Length 141;
Best Local Similarity 87.5%; Pred. No. 4;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 LDRFLATV 9
Db      125 LDKFLATV 132

RESULT 4
HARTNG
hemoglobin alpha chain - northern gundi
C/Species: Ctenodactylus gundi (northern gundi)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 12-Jul-2004
C/Accession: S13282; B33082
R/Beintema, J.J.; Rodewald, K.; Braunitzer, G.
Biol. Chem. Hoppe-Seyler 371, 1089-1099, 1990
A/Title: The primary structures of gundi (Ctenodactylus gundi, Rodentia) hemoglobin and
A/Reference number: S13282; MUID:91197427; PMID:2085415
A/Accession: S13282
A/Molecule type: protein
A/Residues: 1-141 <BRI>
A/Cross-references: UNIPROT:P20854
C/Superfamily: globin; globin homology
C/Keywords: blood; chromoprotein; erythrocyte; heme; heterotetramer; iron; metalloprotein
F/2-141/Domain: globin homology <GLB>
F/58/Binding site: oxygen (His) (distal axial ligand) #status predicted
F/87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match      83.3%; Score 35; DB 1; Length 141;
Best Local Similarity 87.5%; Pred. No. 4;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 LDRFLATV 9
Db      125 LDKFLATV 132

RESULT 5
C25929
hemoglobin alpha-I chain - Kenyan clawed frog
C/Species: Xenopus borealis (Kenyan clawed frog)
C/Date: 25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change 12-Jul-2004
C/Accession: C25929
R/Knoechel, W.; Korge, E.; Basner, A.; Meyerhof, W.
J. Mol. Evol. 23, 211-223, 1986
A/Title: Globin evolution in the genus Xenopus: comparative analysis of cDNAs coding for

```

```

A/Reference number: A92964; MUID:87112760; PMID:3100812
A/Accession: C25929
A/Molecule type: mRNA
A/Residues: 1-141 <KNO>
A/Cross-references: UNIPROT:P07430; GB:M32453; NID:g213941; PIDN:AAA49643.1; PID:g213942
C/Superfamily: globin; globin homology
C/Keywords: chromoprotein; heme; iron; metalloprotein; oxygen carrier
F/2-141/Domain: globin homology <GLB>
F/58/Binding site: oxygen (His) (distal axial ligand) #status predicted
F/87/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match      83.3%; Score 35; DB 2; Length 141;
Best Local Similarity 87.5%; Pred. No. 4;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 LDRFLATV 9
Db      125 LDKFLATV 132

RESULT 6
HAXLI
hemoglobin alpha chain, major - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 12-Jul-2004
C/Accession: A93456; A92863; A90964; S02731; I51428; A02341
R/Kay, R.M.; Harris, R.; Patient, R.K.; Williams, J.G.
Nucleic Acids Res. 11, 1537-1542, 1983
A/Title: Complete nucleotide sequence of a cloned cDNA derived from the major adult alpha
A/Reference number: A93456; MUID:83143364; PMID:6298747
A/Accession: A93456
A/Molecule type: mRNA
A/Residues: 1-142 <KAY>
A/Cross-references: UNIPROT:P02012
R/Partington, G.A.; Baralle, F.E.
J. Mol. Biol. 145, 463-470, 1981
A/Title: Isolation of a Xenopus laevis alpha-globin gene.
A/Reference number: A92863; MUID:81267326; PMID:7265209
A/Accession: A92863
A/Molecule type: DNA
A/Residues: 'S', 30-45, 'N', 47-100; 127-142 <PAR>
R/Richardson, C.; Cappello, J.; Cochran, M.D.; Armentrout, R.W.; Brown, R.D.
Dev. Biol. 78, 161-172, 1980
A/Title: Partial sequence analysis of Xenopus alpha- and beta-globin mRNA as determined
A/Reference number: A90964; MUID:80247171; PMID:6249685
A/Accession: A90964
A/Molecule type: mRNA
A/Residues: 'KUH', 72-96, 'S', 98-113, 'L', 115-142 <RIC>
A/Cross-references: GB:M15381; NID:g213995; PIDN:AAA49662.1; PID:g213996
R/Stalder, J.; Wirthmüller, U.; Beck, J.; Gruber, A.; Meyerhof, W.; Knoechel, W.; Weber
J. Mol. Evol. 28, 64-71, 1988
A/Title: Primary structure and evolutionary relationship between the adult alpha-globin
A/Reference number: S02731; MUID:89178739; PMID:3148743
A/Accession: S02731
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-142 <STA>
A/Cross-references: EMBL:X14259; NID:g64722; PIDN:CAA32472.1; PID:g64723
R/Kay, R.M.; Harris, R.; Patient, R.K.; Williams, J.G.
Nucleic Acids Res. 8, 2691-2707, 1980
A/Title: Molecular cloning of cDNA sequences coding for the major alpha- and beta-globin
A/Reference number: I51428; MUID:81053706; PMID:7001356
A/Accession: I51428
A/Status: translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 129-142 <KAZ>
A/Cross-references: GB:M10577; NID:g214201; PIDN:AAA49732.1; PID:g214202
C/Genetics:
A/Introns: 32/2; 100/3
C/Superfamily: globin; globin homology
C/Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrier
F/3-142/Domain: globin homology <GLB>
F/59/Binding site: oxygen (His) (distal axial ligand) #status predicted

```

F;88/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 83.3%; Score 35; DB 1; Length 142;  
Best Local Similarity 87.5%; Pred. No. 4;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDRFLATV 9  
||:|||||  
Db 126 LDKFLATV 133

RESULT 7  
HAXL2  
hemoglobin alpha chain, minor - African clawed frog  
N;Alternate names: hemoglobin alpha-II chain  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 12-Jul-2004  
A;Accession: A02342; S02733  
R;Knoechel, W.; Meyerhof, W.; Hummel, S.; Grundmann, U.  
Nucleic Acids Res. 11, 1543-1553, 1983  
A;Title: Molecular cloning and sequencing of mRNAs coding for minor adult globin polypep  
A;Reference number: A93457; MUID:83143365; PMID:6298748  
A;Accession: A02342  
A;Molecule type: mRNA  
A;Residues: 1-142 <KNO>  
A;Cross-references: UNIPROT:P02013; GB:X01559; NID:G64536; PIDN:CAA25712.1; PID:G64537  
R;Scauder, J.; Wirthmuller, U.; Beck, J.; Gruber, A.; Meyerhof, W.; Knoechel, W.; Weber  
J. Mol. Evol. 28, 64-71, 1988  
A;Title: Primary structure and evolutionary relationship between the adult alpha-globin  
A;Reference number: S02731; MUID:89178739; PMID:3148743  
A;Accession: S02733  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-23, 'T', 25-29, 'L', 31-142 <STA>  
A;Cross-references: EMBL:X14261; NID:G64724; PIDN:CAA32474.1; PID:G64725  
C;Genetics:  
A;Introns: 32/2; 100/3  
C;Superfamily: Globin; globin homology  
C;Keywords: blood; chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrie  
F;3-142/Domain: globin homology <GLB>  
F;59/Binding site: oxygen (His) (distal axial ligand) #status predicted  
F;88/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 83.3%; Score 35; DB 1; Length 142;  
Best Local Similarity 87.5%; Pred. No. 4;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDRFLATV 9  
||:|||||  
Db 126 LDKFLATV 133

RESULT 8  
JE0264  
cyclin I - human  
C;Species: Homo sapiens (man)  
C;Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: JE0264  
R;Zhu, X.; Naz, R. K.  
Biochem. Biophys. Res. Commun. 249, 56-60, 1998  
A;Title: Expression of a novel isoform of cyclin I in human testis.  
A;Reference number: JE0264; MUID:99381026; PMID:9705831  
A;Accession: JE0264  
A;Molecule type: mRNA  
A;Residues: 1-178 <ZHU>  
A;Cross-references: UNIPROT:Q14094  
C;Comment: This protein may have a physiological role in spermatogenesis and/or human sp

Query Match 83.3%; Score 35; DB 2; Length 178;  
Best Local Similarity 88.9%; Pred. No. 5.1;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLDRFLATV 9

Db 72 LLDGFLATV 80  
||| |||||  
||| |||||

RESULT 9  
S73688  
hypothetical protein H11366 - Mycoplasma pneumoniae (strain ATCC 29342)  
N;Alternate names: hypothetical protein P01\_orf197  
C;Species: Mycoplasma pneumoniae  
A;Variety: ATCC 29342  
C;Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C;Accession: S73688  
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B. C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A;Reference number: S73327; MUID:97105885; PMID:8948633  
A;Accession: S73688  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-197 <HIM>  
A;Cross-references: UNIPROT:P75305; EMBL:AE000034; GB:U00089; NID:G1674031; PIDN:AAB9601  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C;Genetics:  
A;Genetic code: SGC3  
C;Superfamily: acyl carrier protein phosphodiesterase

Query Match 81.0%; Score 34; DB 2; Length 197;  
Best Local Similarity 87.5%; Pred. No. 9.4;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLDRFLAT 8  
||:|||||  
Db 21 LLERFLAT 28

RESULT 10  
AG3446  
arsenate reductase (BC l...-) [imported] - Brucella melitensis (strain 16M)  
C;Species: Brucella melitensis  
C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C;Accession: AG3446  
R;DelVecchio, V. G.; Kaputal, V.; Redkar, R. J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P. H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
A;Reference number: AD3252; PMID:11756688  
A;Accession: AG3446  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-116 <KUR>  
A;Cross-references: UNIPROT:Q8YFG5; UNIPROT:Q8G2F3; GB:AE008917; PIDN:AALS2738.1; PID:gi  
A;Experimental source: strain 16M  
C;Genetics:  
A;Gene: BMEI1557  
A;Map position: I  
C;Keywords: oxidoreductase

Query Match 78.6%; Score 33; DB 2; Length 116;  
Best Local Similarity 87.5%; Pred. No. 8.6;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LDRFLATV 9  
||| |||||  
Db 43 LDRFLKTV 50

RESULT 11  
JC4579  
cyclin D2 - chicken  
N;Alternate names: Cyl D2  
C;Species: Gallus gallus (chicken)  
C;Date: 08-Mar-1996 #sequence\_revision 19-Apr-1996 #text\_change 12-Jul-2004  
C;Accession: JC4579

R;Li, H.; Grenet, J.; Kidd, V.J.

Gene 167, 341-342, 1995

A;Title: Structure and gene expression of avian cyclin D2.

A;Reference number: JC4579; MUID:96144302; PMID:8566807

A;Accession: JC4579

A;Molecule type: mRNA

A;Residues: 1-291 <LH>

A;Cross-references: UNIPROT:P49706; GB:U28980; NID:9968968; PIDN:AAA96955.1; PID:9968968

A;Experimental source: U99 T-cell

C;Comment: This protein acts as a regulator of the cell-cycle-dependent protein kinase 99 product, binds to these products in a combinatorial fashion, activates their enzyme activity

C;Genetics:

A;Gene: cycl2

C;Superfamily: cyclin, A/B/D/E type

C;Keywords: cell cycle control

F;29-193/Region: cyclin-box similarity

Query Match 78.6%; Score 33; DB 2; Length 291;  
Best Local Similarity 87.5%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LDRFLATV 9

Db 84 LDRELAUV 91

|||||

RESULT 12

B75496

moXR protein - Deinococcus radiodurans (strain R1)

C;Species: Deinococcus radiodurans

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C;Accession: B75496

R;White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: B75496

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-320 <WHI>

A;Cross-references: UNIPROT:Q9RWP5; GB:AE001920; GB:AE000513; NID:G6458318; PIDN:AAF1020

A;Experimental source: strain R1

C;Genetics:

A;Map position: 1

C;Superfamily: methanol dehydrogenase regulatory protein

Query Match 78.6%; Score 33; DB 2; Length 320;

Best Local Similarity 87.5%; Pred. No. 26;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LDRFLATV 9

Db 177 LDRELMTV 184

|||||

RESULT 13

A81339

gcpe protein homolog Cj0686 [imported] - Campylobacter jejuni (strain NCTC 11168)

C;Species: Campylobacter jejuni

C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004

C;Accession: A81339

R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel

Nature 403, 665-668, 2000

A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A;Reference number: A81250; MUID:20150912; PMID:10688204

A;Accession: A81339

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-357 <PAR>

A;Cross-references: UNIPROT:Q9PPM1; GB:AL139076; GB:AL111168; NID:96968128; PIDN:CAB72961

A;Experimental source: serotype O2, strain NCTC 11168

C;Genetics:

A;Gene: gcpe; Cj0686

C;Superfamily: gcpe protein

Query Match 78.6%; Score 33; DB 2; Length 357;

Best Local Similarity 77.8%; Pred. No. 29;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LDRFLATV 9

Db 338 LVDRFLAEV 346

|||||

RESULT 14

JC7080

melanin-concentrating hormone receptor [validated] - human

N;Alternate names: MCHR; orphan somatostatin-like receptor 1 (SLC-1)

C;Species: Homo sapiens (man)

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C;Accession: JC7080

R;Shimamura, Y.; Mori, M.; Sugo, T.; Ishibashi, Y.; Abe, M.; Kurokawa, T.; Onda, H.; Nish

Biochem. Biophys. Res. Commun. 261, 622-626, 1999

A;Title: Isolation and identification of melanin-concentrating hormone as the endogenous

A;Reference number: JC7080; MUID:99373129; PMID:10441476

A;Accession: JC7080

A;Molecule type: mRNA

A;Residues: 1-422 <SHI>

A;Cross-references: UNIPROT:Q99705

A;Note: it is uncertain whether Met-1, Met-6 or Met-70 is the initiation codon

C;Superfamily: neurokinin 1 receptor

C;Keywords: hormone receptor; transmembrane protein

Query Match 78.6%; Score 33; DB 2; Length 422;

Best Local Similarity 75.0%; Pred. No. 35;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDRFLATV 9

Db 208 IDRYLATV 215

|||||

RESULT 15

JC7677

allatostatin receptor - American cockroach

C;Species: Periplaneta americana (American cockroach)

C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C;Accession: JC7677

R;Auerwald, L.; Birguel, N.; Gaede, G.; Kreienkamp, H.J.; Richter, D.

Biochem. Biophys. Res. Commun. 282, 904-909, 2001

A;Title: Structural, functional, and evolutionary characterization of novel members of t

A;Reference number: JC7677; MUID:21250673; PMID:11352636

A;Accession: JC7677

A;Molecule type: mRNA

A;Residues: 1-423 <AUE>

A;Cross-references: UNIPROT:Q964D4; GB:AF336364

C;Comment: This receptor, for the neuropeptide, is involved in the inhibition of juvenil

C;Genetics:

A;Gene: Alstr

C;Keywords: neuropeptide; transmembrane protein

F;73-94/Domain: transmembrane #status predicted <TM1>

F;105-127/Domain: transmembrane #status predicted <TM2>

F;143-164/Domain: transmembrane #status predicted <TM3>

F;165-167/Region: active element DRF

F;187-207/Domain: transmembrane #status predicted <TM4>

F;243-264/Domain: transmembrane #status predicted <TM5>

F;290-312/Domain: transmembrane #status predicted <TM6>

F;327-350/Domain: transmembrane #status predicted <TM7>

Query Match 78.6%; Score 33; DB 2; Length 423;

Best Local Similarity 87.5%; Pred. No. 35;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 LDRFLATV 9  
| | | | |  
Db 164 LDRFLAVV 171

Search completed: November 18, 2005, 01:09:40  
Job time : 11.0032 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:04:26 ; Search time 39.9677 Seconds  
(without alignments)  
115.311 Million cell updates/sec

Title: US-10-006-177-8

Perfect score: 42

Sequence: 1 LDRFLATV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	377	1 CYCI_HUMAN	Q14094
2	42	100.0	377	1 CYCI_MOUSE	Q922V9
3	42	100.0	377	2 Q8C7E2	Q922V9
4	42	100.0	377	2 Q8C7E2	Q922V9
5	40	95.2	382	2 Q6P7H3	Q922V9
6	38	90.5	397	2 Q82ST8	Q922V9
7	37	88.1	881	2 Q9A7D2	Q922V9
8	36	85.7	257	2 Q8GD76	Q922V9
9	36	85.7	355	2 Q6NUZ7	Q922V9
10	36	85.7	355	2 Q803P0	Q922V9
11	36	85.7	645	2 Q72R38	Q922V9
12	36	85.7	645	2 Q8F468	Q922V9
13	36	85.7	3121	2 Q42269	Q922V9
14	36	85.7	3148	1 HD_FUGRU	Q922V9
15	35	83.3	141	1 HBAI_XENBO	Q922V9
16	35	83.3	141	1 HBAI_XENLA	Q922V9
17	35	83.3	141	1 HBA2_XENLA	Q922V9
18	35	83.3	141	1 HBA2_CTEGU	Q922V9
19	35	83.3	141	1 HBA2_CTEGU	Q922V9
20	35	83.3	141	1 HBA2_CTEGU	Q922V9
21	35	83.3	141	1 HBA2_CTEGU	Q922V9
22	35	83.3	141	1 HBA2_CTEGU	Q922V9
23	35	83.3	141	1 HBA2_CTEGU	Q922V9
24	35	83.3	141	1 HBA2_CTEGU	Q922V9
25	35	83.3	141	1 HBA2_CTEGU	Q922V9
26	35	83.3	141	1 HBA2_CTEGU	Q922V9
27	35	83.3	141	1 HBA2_CTEGU	Q922V9
28	35	83.3	141	1 HBA2_CTEGU	Q922V9
29	34	81.0	109	2 Q8PR10	Q922V9
30	34	81.0	184	2 Q6YR71	Q922V9
31	34	81.0	197	1 ACPD_MYCPN	Q922V9
32	34	81.0	251	2 Q64XE3	Q922V9

32	34	81.0	299	2	Q6NYI9	Q6ny19 brachydanio
33	34	81.0	299	2	Q6TEN3	Q6ten3 brachydanio
34	34	81.0	299	2	Q7TX1	Q7tx1 brachydanio
35	34	81.0	446	2	Q88PR9	Q88pr9 pseudomonas
36	34	81.0	794	2	Q7S9C2	Q7s9c2 neurospora
37	34	81.0	989	2	Q8NUJ8	Q8nuj8 pichia past
38	34	81.0	990	1	K6P1_PICPA	Q92448 pichia past
39	34	81.0	3311	2	Q7N2F0	Q7n2f0 photorhabdu
40	34	81.0	5216	2	Q7N239	Q7n239 photorhabdu
41	34	81.0	5457	2	Q7N2F7	Q7n2f7 photorhabdu
42	34	81.0	16367	2	Q7N3P5	Q7n3p5 photorhabdu
43	33	78.6	116	2	Q8YFG5	Q8yfg5 bruceella me
44	33	78.6	116	2	Q8G2F3	Q8g2f3 bruceella su
45	33	78.6	190	2	O57481	O57481 stizostedio

#### ALIGNMENTS

#### RESULT 1

ID	CYCI_HUMAN	STANDARD;	PRT;	377 AA.
AC	Q14094;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Cyclin I.			
GN	Name=CCNI;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Forebrain cortex;			
RX	MEDLINE=96086776; PubMed=7493655; DOI=10.1006/excr.1995.1406;			
RA	Nakamura T., Sanokawa R., Sasaki Y.F., Ayusawa D., Oishi M., Mori N.;			
RT	"Cyclin I: a new cyclin encoded by a gene isolated from human brain.";			
RL	Exp. Cell Res. 221:534-542(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Chen J.H., Luo W.Q., Zhou Y., Zhou H.J., Huang X.W., Yuan J.G.,			
RA	Qiang B.Q.;			
RT	"Isolating a new cDNA coding for human cyclin protein.";			
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANT ILE-207.			
RA	Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,			
RA	Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,			
RA	Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;			
RT	"NIH-SNPs, environmental genome project, NIHES ES15478, Department			
RL	of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";			
RN	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.			
RP	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney, and Muscle;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Straube R.L., Reingold E.A., Grouse L.H., Derge J.G.			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP SEQUENCE OF 1-178 FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=98381026; PubMed=9705831; DOI=10.1006/bbrc.1998.9052;  
 RA Zhu X., Naz R.K.;  
 RT "Expression of a novel isoform of cyclin I in human testis.";   
 RL Biochem. Biophys. Res. Commun. 249:56-60(1998).  
 CC -1- TISSUE SPECIFICITY: Highest levels in adult heart, brain and  
 CC skeletal muscle. Lower levels in adult placenta, lung, kidney and  
 CC pancreas. Also high levels in fetal brain and lower levels in  
 CC fetal lung, liver and kidney. Also abundant in testis and thyroid.  
 CC -1- DEVELOPMENTAL STAGE: Expression is independent of the cell cycle  
 CC in lung fibroblasts.  
 CC -1- SIMILARITY: Belongs to the cyclin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D50310; BAA08849.1; -;  
 DR EMBL; AY207372; AA013492.1; -;  
 DR EMBL; BC000420; AAH00420.1; -;  
 DR EMBL; BC004975; AAH04975.1; -;  
 DR EMBL; AF135162; AAF43786.1; -;  
 DR PIR; JEO264; JEO264.  
 DR Genew; HGNC:1595; CCNI.  
 DR H-invDB; HIX0004313; -;  
 DR GO; GO:0007283; P:spermatogenesis; NAS.  
 DR InterPro; IPR006670; Cyclin.  
 DR InterPro; IPR011028; Cyclin-like.  
 DR Pfam; PF00134; Cyclin\_N; 1.  
 DR SMART; SM00385; CYCLIN; 1.  
 DR PROSITE; PS00292; CYCLINS; FALSE\_NEG.  
 KW Cyclin; Polymorphism.  
 FT VARIANT 207 207 V -> I (in dbSNP:4252903).  
 FT CONFLICT 9 9 N -> D (in Ref. 5).  
 FT CONFLICT 58 58 Q -> R (in Ref. 5).  
 FT CONFLICT 75 75 R -> G (in Ref. 5).  
 SQ SEQUENCE 377 AA; 42557 MW; 2DE84EFA74698F6C CRC64;  
 Query Match 100.0%; Score 42; DB 1; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 5.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 LLDRFLATV 9  
 Db 72 LLDRFLATV 80  
 RESULT 2  
 ID CYCI MOUSE STANDARD; PRT; 377 AA.  
 AC Q922V9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Cyclin I.  
 GN Name=Ccni;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=99172101; PubMed=10072591;  
 RA Jensen M.R., Audolfsson T., Keck C.L., Zimonjic D.B.,  
 RA Thorgerirsson S.S.;  
 RT "Assignment of the cyclin I gene (Ccni) to mouse chromosome 5E3.3-F1.  
 RT 3 by in situ hybridization.";   
 RL Cytogenet. Cell Genet. 83:242-243(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20509992; PubMed=11054536; DOI=10.1016/S0378-1119(00)00361-9;  
 RA Jensen M.R., Audolfsson T., Factor V.W., Thorgerirsson S.S.;  
 RT "In vivo expression and genomic organization of the mouse cyclin I  
 RT gene (Ccni).";   
 RL Gene 256:59-67(2000).  
 CC -1- SIMILARITY: Belongs to the cyclin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF005886; AAD01253.2; -;  
 DR EMBL; AF228740; AAF43391.1; -;  
 DR EMBL; AF228739; AAF43391.1; JOINED.  
 DR MGD; MGI:1341077; Ccni.  
 DR InterPro; IPR006670; Cyclin.  
 DR InterPro; IPR011028; Cyclin-like.  
 DR Pfam; PF00134; Cyclin\_N; 1.  
 DR SMART; SM00385; CYCLIN; 1.  
 DR PROSITE; PS00292; CYCLINS; FALSE\_NEG.  
 KW Cyclin.  
 SQ SEQUENCE 377 AA; 42261 MW; 4151141C6D9AE677 CRC64;  
 Query Match 100.0%; Score 42; DB 1; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 5.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 LLDRFLATV 9  
 Db 72 LLDRFLATV 80  
 RESULT 3  
 ID Q8C7E2 PRELIMINARY; PRT; 377 AA.  
 AC Q8C7E2;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Mus musculus adult pancreas islet cells cDNA, RIKEN full-length  
 DE enriched library, clone:CS20001G04 product:cyclin I, full insert  
 DE sequence.  
 GN Name=Ccni;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";   
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";
 RN Nature 409:685-690(2001).
 RP [3]
 RC SEQUENCE FROM N.A.
 RA STRAIN=C57BL/6J; TISSUE=Pancreas;
 RA The RIKEN Consortium;
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RA "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RA "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multipipette sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiracka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Takagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
 RA Tonaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the cyclin family.
 DR EMBL; AK050465; BAC34271.1; -.
 DR MGD; MGI:1341077; Ccni.
 DR GO; GO:0000074; P:regulation of cell cycle; TAS.
 DR InterPro; IPR006670; Cyclin.
 DR InterPro; IPR011028; Cyclin like.
 DR InterPro; IPR006671; Cyclin\_N.
 DR Pfam; PF00134; Cyclin\_N; 1.
 DR SMART; SM00385; CYCLIN; 1.
 KW Cyclin.
 SQ SEQUENCE 377 AA; 42206 MW; A1B803EBE135B0E8 CRC64;
 Query Match 100.0%; Score 42; DB 2; Length 377;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 LLDRFLATV 9
 Db 72 LLDRFLATV 80
 RESULT 4
 Q99LF2 PRELIMINARY; PRT; 377 AA.
 AC Q99LF2;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cyclin I.
 GN Name=Ccni;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI\_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Alschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the cyclin family.
 DR EMBL; BC003290; AAH03290.1; -.
 DR MGD; MGI:1341077; Ccni.
 DR GO; GO:0000074; P:regulation of cell cycle; TAS.
 DR InterPro; IPR006670; Cyclin.
 DR InterPro; IPR011028; Cyclin like.
 DR InterPro; IPR006671; Cyclin\_N.
 DR Pfam; PF00134; Cyclin\_N; 1.
 DR SMART; SM00385; CYCLIN; 1.
 KW Cyclin.
 SQ SEQUENCE 377 AA; 42227 MW; FDA2D896A5366A9F CRC64;
 Query Match 100.0%; Score 42; DB 2; Length 377;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 LLDRFLATV 9
 Db 72 LLDRFLATV 80
 RESULT 5
 Q6P7H3 PRELIMINARY; PRT; 382 AA.
 ID Q6P7H3
 AC Q6P7H3;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE MGC68660 protein.
 GN Name=MGC68660;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI\_TaxID=8355;

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RX  MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA  Datschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA  Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA  Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA  Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA  Jones S.J., Marra W.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
RL  and mouse cDNA sequences.";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RX  MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA  Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA  Richardson P.;
RT  "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL  Dev. Dyn. 225:384-391(2002).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RA  Klein S., Strausberg R.;
RL  Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC  -I- SIMILARITY: Belongs to the cyclin family.
DR  GO; GO:0061670; AAHE1670.1; -.
DR  GO; GO:0007049; P:cell cycle; IEA.
DR  GO; GO:0000910; P:cytokinesis; IEA.
DR  InterPro; IPR006670; Cyclin.
DR  InterPro; IPR011028; Cyclin like.
DR  InterPro; IPR006671; Cyclin_N.
DR  Pfam; PF00134; Cyclin N; 1.
DR  SMART; SM00385; CYCLIN; 1.
DR  PROSITE; PS00292; CYCLINS; 1.
KW  Cell cycle; Cell division; Cyclin.
SQ  SEQUENCE 382 AA; 42913 MW; E19D7C5D8FEC626A CRC64;

Query Match 95.2%; Score 40; DB 2; Length 382;
Best Local Similarity 88.9%; Pred. NO. 15;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9
Db 74 ILDRFLATV 82

RESULT 6
Q82ST8 PRELIMINARY; PRT; 397 AA.
AC Q82ST8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FAD-dependent pyridine nucleotide-diaphorase oxidoreductase (EC
DE 1.6.99.3).
GN OrderedLocusNames=NF2216;
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.

```

```

OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IPO 14298;
RX MEDLINE=22586410; PubMed=12700255;
RX DOI=10.1128/JB.185.9.2759-2773.2003;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.L.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Savavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RL obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773(2003).
DR EMBL; BX321863; CAD86128.1; -.
DR GO; GO:0015036; F:disulfide oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001327; FAD_Pyr_redox.
DR InterPro; IPR000103; Pyridine_redox_2.
DR Pfam; PF00070; Pyr_redox; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00469; PNDRTASEII.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 397 AA; 44141 MW; E38B5CAE8031D34A CRC64;

Query Match 90.5%; Score 38; DB 2; Length 397;
Best Local Similarity 100.0%; Pred. NO. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLAT 8
Db 154 LLDRFLAT 161

RESULT 7
Q9A7D2 PRELIMINARY; PRT; 881 AA.
AC Q9A7D2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE TonB-dependent receptor.
GN OrderedLocusNames=CC1791;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RC MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Piscocha I., Nelson W.C., Newton A., Stephens C., Phacke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005853; AAK23767.1; -.
DR PIR; C87471; C87471.
DR TIGR; CC1791; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR010104; TonB_receptorbac.
DR TIGRFAMs; TIGR01782; TonB-Xanth-Caul; 1.
KW Complete proteome; Receptor.
SQ SEQUENCE 881 AA; 97124 MW; 7492D0758023D790 CRC64;

Query Match 88.1%; Score 37; DB 2; Length 881;
Best Local Similarity 77.8%; Pred. NO. 1.3e+02;

```

XX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalska U., Smalls D.E., Schmerch A., Schein J.E.,
RA	Jones S.J., Maier M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences."
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Embryo;
RA	Director MGC Project;
RA	Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC	-I- SIMILARITY: Belongs to the cyclin family.
DR	EMBL: BC068369; AAH68369.1; -
DR	InterPro: IPR006670; Cyclin.
DR	InterPro: IPR011028; Cyclin_like.
DR	InterPro: IPR006671; Cyclin_N.
DR	Pfam: PF00134; Cyclin_N.1.
DR	SMART: SM00385; CYCLIN; 1.
DR	PROSITE: PS00292; CYCLINS; UNKNOWN_1.
CY	Cyclin.
KW	SEQUENCE 355 AA; 39614 MW; B9D4236191F663BF CRC64;
SQ	
	Query Match 85.7%; Score 36; DB 2; Length 355;
	Best Local Similarity 66.7%; Pred. No. 87;
	Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0
Qy	1 LLDRFLATV 9
	:     :
Db	74 ILDRFLSTI 82
RESULT 10	
Q803P0	PRELIMINARY; PRt; 355 AA.
ID	
AC	Q803P0;
DT	01-JUN-2003 (TrEMBLrel. 24, Created)
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Cyclin I.
DE	Cyclin I.
GN	Names=ccni;
OS	Brachydanio rerio (zebrafish) (Danio rerio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC	Cyprinidae; Danio.
OX	NCBI_TaxID=7955;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	SPRAIN=AB; TISSUE-whole body;
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalska U., Smalls D.E., Schmerch A., Schein J.E.,
RA	Jones S.J., Maier M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences."
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Embryo;
RA	Director MGC Project;
RA	Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC	-I- SIMILARITY: Belongs to the cyclin family.
DR	EMBL: BC068369; AAH68369.1; -
DR	InterPro: IPR006670; Cyclin.
DR	InterPro: IPR011028; Cyclin_like.
DR	InterPro: IPR006671; Cyclin_N.
DR	Pfam: PF00134; Cyclin_N.1.
DR	SMART: SM00385; CYCLIN; 1.
DR	PROSITE: PS00292; CYCLINS; UNKNOWN_1.
CY	Cyclin.
KW	SEQUENCE 355 AA; 39614 MW; B9D4236191F663BF CRC64;
SQ	
	Query Match 85.7%; Score 36; DB 2; Length 355;
	Best Local Similarity 66.7%; Pred. No. 87;
	Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
RA Fahney J., Helton E., Kettelman M., Madan A.C., Shevchenko Y., Bouffard G.G.,  
RA Whiting J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN (2)  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB; TISSUE=whole body;  
RA Strauberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to the cyclin family.  
DR EMBL; BC044400; AAH44400.1; -;  
DR HSSP; P30274; 1VIN  
DR ZFIN; ZDB-GENE-040426-2898; ceni.  
DR InterPro; IPR006670; Cyclin.  
DR InterPro; IPR011028; Cyclin\_like.  
DR InterPro; IPR006671; Cyclin\_N.  
DR Pfam; PF00134; Schmutz J.,  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; UNKNOWN\_1.  
KW Cyclin.  
SQ SEQUENCE 355 AA; 39517 MW; D29999F2E9812C880 CRC64;

Query Match 85.7%; Score 36; DB 2; Length 355;  
Best Local Similarity 66.7%; Pred. No. 87;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9  
:|||||:  
Db 74 ILDRFLSTI 82

RESULT 11  
Q72RJ8  
ID Q72RJ8 PRELIMINARY; PRT; 645 AA.  
AC Q72RJ8  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Long-chain-fatty-acid CoA ligase.  
GN Name=ydid; OrderedLocusNames=LIC11747;  
OS Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar  
OS Copenhagen)  
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiaceae; Leptospira.  
OX NCBI\_TaxID=44275;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=Flotruz L1-130;  
RX PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;  
RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Montecarlo C.B.,  
RA Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,  
RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrer H.,  
RA Coutinho L.L., Degraive W.M., Dellagostin O.A., El-Dorri H.,  
RA Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Gigliotti E.A.,  
RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,  
RA Jeronimo S.M.B., Junqueira-de-Azevedo I.L.M., Kimura E.T.,  
RA Kurame E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,  
RA de Oliveira R.C., Pereira G.G., Reis M.S., Schriefer A.,  
RA Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,  
RA Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;  
RT "Comparative genomics of two Leptospira interrogans serovars reveals  
RT novel insights into physiology and pathogenesis."  
RL J. Bacteriol. 186:2164-2172(2004).  
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme  
CC family.  
DR EMBL; AE017294; AAS70336.1; -;  
DR GO; GO:0016874; F:ligase activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR000873; AMP-bind.  
DR PRINTS; PR00154; AMPBINDING.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
KW Complete proteome; Ligase.  
SQ SEQUENCE 645 AA; 72542 MW; 96F64565E5B952AC CRC64;

Query Match 85.7%; Score 36; DB 2; Length 645;  
Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9  
:|||||:  
Db 341 LLDRFIASV 349

RESULT 12  
Q8F468  
ID Q8F468 PRELIMINARY; PRT; 645 AA.  
AC Q8F468  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Long-chain-fatty-acid CoA ligase (EC 6.2.1.3).  
GN OrderedLocusNames=LA2177;  
OS Leptospira interrogans.  
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiaceae; Leptospira.  
OX NCBI\_TaxID=173;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
RX MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;  
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Mao Y.-G., Xu H.,  
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,  
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,  
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,  
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,  
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,  
RA Xu J.-G., Zhao G.-P.;  
RT "Unique physiological and pathogenic features of Leptospira  
RT interrogans revealed by whole-genome sequencing."  
RL Nature 422:888-893(2003).  
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme  
CC family.  
DR EMBL; AE011388; AAM49376.1; -;  
DR HSSP; P08659; ILIC1.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR GO; GO:0004467; F:long-chain-fatty-acid-CoA ligase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR000873; AMP-bind.  
DR Pfam; PF00501; AMP-binding; 2.  
DR PRINTS; PR00154; AMPBINDING.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
KW Complete proteome.  
SQ SEQUENCE 645 AA; 72544 MW; E9DPE2B6A1718911 CRC64;

Query Match 85.7%; Score 36; DB 2; Length 645;  
Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9  
:|||||:  
Db 341 LLDRFIASV 349

RESULT 13  
ID O42269 PRELIMINARY; PRT; 3121 AA.  
AC O42269  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Huntingtin.

GN Name=hd;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90014243; PubMed=9795179; DOI=10.1016/S0378-1119(98)00342-4;  
 RA Karlovich C.A., John R.M., Ramirez L., Stainier D.Y., Myers R.M.;  
 RT "Characterization of the Huntington's disease (HD) gene homologue in  
 the zebrafish Danio rerio.";  
 RL Gene 217:117-125(1998).  
 DR EMBL; AF052602; AAC63983.1; -.  
 DR ZFIN; ZDB-GENE-990415-131; hd.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR000357; HEAT.  
 DR Pfam; PF02985; HEAT\_5.  
 DR PRINTS; PR00375; HUNTINGTIN.  
 SQ SEQUENCE 3121 AA; 346666 MW; 3CC884E199D324AC CRC64;  
  
 Query Match 85.7%; Score 36; DB 2; Length 3121;  
 Best Local Similarity 88.9%; Pred. No. 7.5e+02;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 Qy 1 LLDRFLATV 9  
 Db 2037 LLDRFRATV 2045  
 ||||| |||  
  
 RESULT 14  
 HD\_FUGRU  
 ID HD\_FUGRU STANDARD; PRT; 3148 AA.  
 AC P51112;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Huntington (Huntington's disease protein homolog) (HD protein).  
 GN Name=HD;  
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontoidea; Tetraodontidae; Takifugu.  
 OX NCBI\_TaxID=31033;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95375788; PubMed=7647794;  
 RA Baxendale S., Abdulla S., Elgar G., Buck D., Berks M., Micklem G.,  
 RA Durbin R., Bates G., Brenner S., Beck S., Lehrach H.;  
 RT "Comparative sequence analysis of the human and pufferfish  
 Huntington's disease genes.";  
 RL Nat. Genet. 10:67-76(1995).  
 CC -1- FUNCTION: May play a role in microtubule-mediated transport or  
 vesicle function.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- POLYMORPHISM: The poly-Gln region (four residues) does not appear  
 to be polymorphic, explaining the absence of a HD-like disorder.  
 CC -1- SIMILARITY: Belongs to the huntingtin family.  
 CC -1- SIMILARITY: Contains 10 HEAT repeats.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; X82939; CAA58112.1; -.  
 DR InterPro; IPR008938; ARM.

DR InterPro; IPR000357; HEAT.  
 DR InterPro; IPR000091; Huntingtin.  
 DR Pfam; PF03541; Huntingtin; 1.  
 DR PRINTS; PR00375; HUNTINGTIN.  
 DR PROSITE; PS50077; HEAT\_REPEAT; 1.  
 KW Repeat.  
 FT REPEAT 149 186 HEAT 1.  
 FT REPEAT 191 228 HEAT 2.  
 FT REPEAT 760 797 HEAT 3.  
 FT REPEAT 861 898 HEAT 4.  
 FT REPEAT 1419 1456 HEAT 5.  
 FT DOMAIN 18 21 Poly-Gln.  
 FT DOMAIN 679 682 Poly-Ala.  
 FT DOMAIN 1104 1108 Poly-Ser.  
 SQ SEQUENCE 3148 AA; 348932 MW; D9358676B0345243 CRC64;  
  
 Query Match 85.7%; Score 36; DB 1; Length 3148;  
 Best Local Similarity 88.9%; Pred. No. 7.5e+02;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 Qy 1 LLDRFLATV 9  
 Db 2064 LLDRFRATV 2072  
 ||||| |||  
  
 RESULT 15  
 HBAL\_XENBO  
 ID HBAL\_XENBO STANDARD; PRT; 141 AA.  
 AC P07430;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 29-MAR-2004 (Rel. 43, Last annotation update)  
 DE Hemoglobin alpha-1 chain (Hemoglobin alpha-major chain).  
 OS Xenopus borealis (Kenyan clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8354;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87112760; PubMed=3100812;  
 RA Knoechel W., Korge E., Baegner A., Meyerhof W.;  
 RT "Globin evolution in the genus Xenopus: comparative analysis of cDNAs  
 coding for adult globin polypeptides of Xenopus borealis and Xenopus  
 tropicalis.";  
 RL J. Mol. Evol. 23:211-223(1986).  
 CC -1- FUNCTION: Involved in oxygen transport from the lung to the  
 various peripheral tissues.  
 CC -1- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.  
 CC -1- TISSUE SPECIFICITY: Red blood cells.  
 CC -1- SIMILARITY: Belongs to the globin family.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; M32453; AAA49643.1; -.  
 DR PIR; C25929; C25929.  
 DR HSP; P01922; 1C7C.  
 DR InterPro; IPR002338; Alpha haem.  
 DR InterPro; IPR000971; Globin.  
 DR InterPro; IPR009050; Globin\_like.  
 DR Pfam; PF00042; Globin; 1.  
 DR PRINTS; PR00612; ALPHAHAEH.  
 DR PROSITE; PS01033; GLOBIN; 1.  
 KW Erythrocyte; Heme; Oxygen transport; Transport.  
 FT INIT MET 0 0  
 FT METAL 58 87 Iron (heme distal ligand).  
 FT METAL 87 87 Iron (heme proximal ligand).

SQ SEQUENCE 141 AA; 15614 MW; 4D55B7E3B080CC95 CRC64;  
 Query Match 83.3%; Score 35; DB 1; Length 141;  
 Best Local Similarity 87.5%; Pred. No. 56;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LDRFLATV 9  
 ||:||||  
 Db 125 LDRFLATV 132

Search completed: November 18, 2005, 01:08:18  
 Job time : 43.9677 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 00:48:52 ; Search time 13.6452 Seconds  
(without alignments)  
49.237 Million cell updates/sec

Title: US-10-006-177-8

Perfect score: 42

Sequence: 1 LLDRFLATV 9

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	254	4	US-09-796-149B-4
2	42	100.0	377	2	US-08-969-106-6
3	42	100.0	377	3	US-09-054-492B-1
4	42	100.0	377	4	US-09-338-125-6
5	35	83.3	120	4	US-09-902-540-14885
6	35	83.3	303	4	US-09-107-532A-5585
7	34	81.0	561	4	US-09-107-532A-5323
8	33	78.6	353	3	US-08-984-288-2
9	33	78.6	353	3	US-09-224-426-4
10	33	78.6	353	3	US-09-478-601-4
11	33	78.6	353	3	US-09-478-602-4
12	33	78.6	353	3	US-09-218-467B-2
13	33	78.6	353	4	US-09-885-478-4
14	33	78.6	353	4	US-09-885-478-28
15	33	78.6	353	4	US-09-919-039-127
16	33	78.6	353	4	US-08-826-509-351
17	33	78.6	353	4	US-08-826-509-355
18	33	78.6	353	4	US-08-826-509-363
19	33	78.6	353	4	US-08-826-509-367
20	33	78.6	353	4	US-08-826-509-371
21	33	78.6	353	4	US-08-826-509-375
22	33	78.6	353	4	US-08-826-509-379
23	33	78.6	353	4	US-08-826-509-383
24	33	78.6	353	4	US-08-826-509-387
25	33	78.6	400	5	PCT-US95-16472-2
26	33	78.6	402	3	US-08-602-809-2
27	33	78.6	402	4	US-08-170-496D-52

28	33	78.6	402	4	US-09-170-496D-192
29	33	78.6	422	3	US-09-224-426-2
30	33	78.6	422	3	US-09-478-601-2
31	33	78.6	422	3	US-09-478-602-2
32	33	78.6	422	4	US-09-885-478-2
33	33	78.6	422	4	US-09-885-478-26
34	33	78.6	422	4	US-09-885-478-27
35	33	78.6	446	4	US-08-826-509-359
36	33	78.6	709	4	US-09-826-509-589
37	32	76.2	554	4	US-09-902-540-15061
38	32	76.2	725	4	US-09-252-991A-24201
39	31	73.8	65	3	US-09-352-078-3
40	31	73.8	82	3	US-08-722-719-60
41	31	73.8	82	4	US-09-334-951-60
42	31	73.8	82	4	US-09-334-923A-60
43	31	73.8	82	4	US-09-334-954A-60
44	31	73.8	82	4	US-09-567-225-5
45	31	73.8	83	3	US-08-722-719-18

#### ALIGNMENTS

RESULT 1  
US-09-796-149B-4  
; Sequence 4, Application US/09796149B  
; Patent No. 6825033  
; GENERAL INFORMATION:  
; APPLICANT: Univ. of Southern California  
; TITLE OF INVENTION: Mutated cyclin G1 protein  
; FILE REFERENCE: 4-31342A/USC  
; CURRENT APPLICATION NUMBER: US/09/796,149B  
; CURRENT FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-796-149B-4

Query Match 100.0%; Score 42; DB 4; Length 254;  
Best Local Similarity 100.0%; Pred. No. 0.82; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Oy 1 LLDRFLATV 9  
Db 33 LLDRFLATV 41  
|||||||

RESULT 2  
US-08-969-106-6  
; Sequence 6, Application US/08969106  
; Patent No. 5986055  
; GENERAL INFORMATION:  
; APPLICANT: Yang, M.  
; APPLICANT: Nandabalan, K.  
; APPLICANT: Schulz, V.  
; TITLE OF INVENTION: CDK2 INTERACTIONS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:

Sequence 192, Appl  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 26, Appl  
Sequence 27, Appl  
Sequence 359, App  
Sequence 589, App  
Sequence 15061, A  
Sequence 24201, A  
Sequence 3, Appli  
Sequence 60, Appl  
Sequence 60, Appl  
Sequence 60, Appl  
Sequence 60, Appl  
Sequence 5, Appli  
Sequence 18, Appl

APPLICATION NUMBER: US/08/969,106  
FILING DATE: 13-NOV-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 7934-057  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 377 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-969-106-6

Query Match 100.0%; Score 42; DB 2; Length 377;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDREFLATV 9  
| | | | | | | | | |  
Db 72 LLDREFLATV 80

RESULT 3  
US-09-054-492B-1  
Sequence 1, Application US/09054492B  
Patent No. 621815  
GENERAL INFORMATION:  
APPLICANT: TAKESHI NAKAMURA  
TITLE OF INVENTION: HUMAN CYCLIN I AND GENES ENCODING SAME  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PILLSBURY, MADISON & SUTRO  
STREET: 1100 NEW YORK AVENUE, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/054,492B  
FILING DATE: APRIL 3, 1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: PAUL E. WHITE, JR.  
REGISTRATION NUMBER: 32,011  
REFERENCE/DOCKET NUMBER: 7898/252159  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627CUSH  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 377  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-054-492B-1

Query Match 100.0%; Score 42; DB 3; Length 377;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDREFLATV 9  
| | | | | | | | | |  
Db 72 LLDREFLATV 80

RESULT 4  
US-09-338-125-6  
Sequence 6, Application US/09338125  
Patent No. 6521412  
GENERAL INFORMATION:  
APPLICANT: Yang, M.  
APPLICANT: Nandabalan, K.  
APPLICANT: Schulz, V.  
TITLE OF INVENTION: CDK2 INTERACTIONS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/338,125  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/969,106  
FILING DATE: 13-NOV-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 7934-057  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 377 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-338-125-6

Query Match 100.0%; Score 42; DB 4; Length 377;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDREFLATV 9  
| | | | | | | | | |  
Db 72 LLDREFLATV 80

RESULT 5  
US-09-902-540-14885  
Sequence 14885, Application US/09902540  
Patent No. 683447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(15849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883  
 ; PRIOR FILING DATE: 2000-07-10  
 ; NUMBER OF SEQ ID NOS: 16825  
 ; SEQ ID NO 14885  
 ; LENGTH: 120  
 ; TYPE: PRT  
 ; ORGANISM: Myxococcus xanthus  
 US-09-902-540-14885

Query Match 83.3%; Score 35; DB 4; Length 120;  
 Best Local Similarity 88.9%; Pred. No. 9.4;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLDREFLATV 9  
 |||||  
 DB 111 LLDGFLATV 119

RESULT 6  
 US-09-107-532A-5585  
 ; Sequence 5585, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD/ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/107,532A  
 ; FILING DATE: 30-Jun-1998  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/085,598  
 ; FILING DATE: 14 May 1998  
 ; APPLICATION NUMBER: 60/051571  
 ; FILING DATE: July 2, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ariniello, Pamela Deneke  
 ; REGISTRATION NUMBER: 40,489  
 ; REFERENCE/DOCKET NUMBER: GTC-012  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (781)893-8277  
 ; TELEFAX: (781)893-8277  
 ; INFORMATION FOR SEQ ID NO: 5585:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 303 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: YES  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Enterococcus faecium  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (B) LOCATION 1...303  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5585:

Query Match 83.3%; Score 35; DB 4; Length 303;  
 Best Local Similarity 77.8%; Pred. No. 26;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLDREFLATV 9  
 |||||  
 DB 111 LLDGFLATV 119

RESULT 7  
 US-09-107-532A-5323  
 ; Sequence 5323, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD/ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/107,532A  
 ; FILING DATE: 30-Jun-1998  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/085,598  
 ; FILING DATE: 14 May 1998  
 ; APPLICATION NUMBER: 60/051571  
 ; FILING DATE: July 2, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ariniello, Pamela Deneke  
 ; REGISTRATION NUMBER: 40,489  
 ; REFERENCE/DOCKET NUMBER: GTC-012  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (781)893-8277  
 ; TELEFAX: (781)893-8277  
 ; INFORMATION FOR SEQ ID NO: 5323:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 561 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: YES  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Enterococcus faecium  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (B) LOCATION 1...561  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5323:

Query Match 81.0%; Score 34; DB 4; Length 561;  
 Best Local Similarity 75.0%; Pred. No. 79;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDRFLATV 9  
 :|||:  
 DB 16 LDRFLATV 23

RESULT 8  
 US-08-984-288-2  
 ; Sequence 2, Application US/08984288  
 ; Patent No. 6033872  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BERGMA, DERK  
 ; APPLICANT: ELLIS, CATHERINE  
 ; TITLE OF INVENTION: NOVEL HUMAN 11CB SPLICE V

Query Match 83.3%; Score 35; DB 4; Length 303;  
 Best Local Similarity 77.8%; Pred. No. 26;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLDREFLATV 9  
 |||||  
 DB 98 LLDREFLATV 106

RESULT 7  
 US-09-107-532A-5323  
 ; Sequence 5323, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD/ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/107,532A  
 ; FILING DATE: 30-Jun-1998  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/085,598  
 ; FILING DATE: 14 May 1998  
 ; APPLICATION NUMBER: 60/051571  
 ; FILING DATE: July 2, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ariniello, Pamela Deneke  
 ; REGISTRATION NUMBER: 40,489  
 ; REFERENCE/DOCKET NUMBER: GTC-012  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (781)893-8277  
 ; TELEFAX: (781)893-8277  
 ; INFORMATION FOR SEQ ID NO: 5323:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 561 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: YES  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Enterococcus faecium  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (B) LOCATION 1...561  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5323:

Query Match 81.0%; Score 34; DB 4; Length 561;  
 Best Local Similarity 75.0%; Pred. No. 79;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDRFLATV 9  
 :|||:  
 DB 16 LDRFLATV 23

RESULT 8  
 US-08-984-288-2  
 ; Sequence 2, Application US/08984288  
 ; Patent No. 6033872  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BERGMA, DERK  
 ; APPLICANT: ELLIS, CATHERINE  
 ; TITLE OF INVENTION: NOVEL HUMAN 11CB SPLICE V

;; TITLE OF INVENTION: ARIANT  
;; NUMBER OF SEQUENCES: 6  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: RAINER & PRESTIA  
;; STREET: P.O. BOX 980  
;; CITY: VALLEY FORGE  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19482  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA: US/08/984,288  
;; APPLICATION NUMBER: US/08/984,288  
;; FILING DATE: 03-DEC-1997  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 60/032,763  
;; FILING DATE: 11-DEC-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: PRESTIA, PAUL F  
;; REGISTRATION NUMBER: 23,031  
;; REFERENCE/DOCKET NUMBER: P50599  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 610-407-0700  
;; TELEFAX: 610-407-0701  
;; TELEX: 846169  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 353 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-984-288-2

Query Match 78.6%; Score 33; DB 3; Length 353;  
Best Local Similarity 75.0%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDRFLATV 9  
Db 139 IDRYLATV 146

RESULT 9  
US-09-224-426-4  
; Sequence 4, Application US/09224426  
; Patent No. 6221613  
; GENERAL INFORMATION:  
; APPLICANT: Salom, John A.  
; APPLICANT: Laz, Thomas M.  
; APPLICANT: Nagorny, Raissa  
; APPLICANT: Wilson, Amy E.  
; TITLE OF INVENTION: DNA Encoding A Human Melanin Concentrating Hormone  
; FILE REFERENCE: 57453/JPW/JHB  
; CURRENT APPLICATION NUMBER: US/09/224,426  
; CURRENT FILING DATE: 1998-12-31  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0 - beta  
; SEQ ID NO 4  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: rat  
US-09-224-426-4

Query Match 78.6%; Score 33; DB 3; Length 353;  
Best Local Similarity 75.0%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDRFLATV 9  
Db 139 IDRYLATV 146

RESULT 10  
US-09-478-601-4  
; Sequence 4, Application US/09478601  
; Patent No. 6221616  
; GENERAL INFORMATION:  
; APPLICANT: Salom, John A.  
; APPLICANT: Laz, Thomas M.  
; APPLICANT: Nagorny, Raissa  
; APPLICANT: Wilson, Amy E.  
; TITLE OF INVENTION: DNA Encoding A Human Melanin Concentrating Hormone  
; FILE REFERENCE: 574532/JPW  
; CURRENT APPLICATION NUMBER: US/09/478,601  
; CURRENT FILING DATE: 2000-01-06  
; EARLIER APPLICATION NUMBER: 09/224,426  
; EARLIER FILING DATE: 1998-12-31  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-478-601-4

Query Match 78.6%; Score 33; DB 3; Length 353;  
Best Local Similarity 75.0%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDRFLATV 9  
Db 139 IDRYLATV 146

RESULT 11  
US-09-478-602-4  
; Sequence 4, Application US/09478602  
; Patent No. 6291195  
; GENERAL INFORMATION:  
; APPLICANT: Salom, John A.  
; APPLICANT: Laz, Thomas M.  
; APPLICANT: Nagorny, Raissa  
; APPLICANT: Wilson, Amy E.  
; TITLE OF INVENTION: DNA Encoding A Human Melanin Concentrating Hormone  
; FILE REFERENCE: 57453v/JPW  
; CURRENT APPLICATION NUMBER: US/09/478,602  
; CURRENT FILING DATE: 2000-01-06  
; EARLIER APPLICATION NUMBER: 09/224,426  
; EARLIER FILING DATE: 1998-12-31  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-478-602-4

Query Match 78.6%; Score 33; DB 3; Length 353;  
Best Local Similarity 75.0%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LDRFLATV 9  
Db 139 IDRYLATV 146

RESULT 12  
US-09-218-467B-2

; Sequence 2, Application US/09218467B  
; Patent No. 6362326  
; GENERAL INFORMATION:  
; APPLICANT: SATHE, GANESH  
; APPLICANT: ELLIS, CATHERINE  
; APPLICANT: HALSEY, WENDY  
; APPLICANT: BERGSMÄ DERK  
; TITLE OF INVENTION: 11cby Genomic Sequence  
; FILE REFERENCE: GP-50010  
; CURRENT APPLICATION NUMBER: US/09/218,467B  
; CURRENT FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-218-467B-2

Query Match 78.6%; Score 33; DB 3; Length 353;  
Best Local Similarity 75.0%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDRFLATV 9  
Db 139 IDRYLATV 146  
:|:|:|

## RESULT 13

US-09-885-478-4  
; Sequence 4, Application US/09885478  
; Patent No. 6723552  
; GENERAL INFORMATION:  
; APPLICANT: SALON, JOHN A  
; APPLICANT: LAZ, THOMAS M  
; APPLICANT: NAGORNY, RAISA  
; APPLICANT: WILSON, AMY E  
; TITLE OF INVENTION: DNA ENCODING A HUMAN MELANIN CONCENTRATING HORMONE RECEPTOR (MCH1  
; FILE REFERENCE: 1795/57453-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/885,478  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: PCT/US99/31169  
; PRIOR FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: RATTUS NORVEGICUS  
US-09-885-478-4

Query Match 78.6%; Score 33; DB 4; Length 353;  
Best Local Similarity 75.0%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDRFLATV 9  
Db 139 IDRYLATV 146  
:|:|:|

## RESULT 14

US-09-885-478-28  
; Sequence 28, Application US/09885478  
; Patent No. 6723552  
; GENERAL INFORMATION:  
; APPLICANT: SALON, JOHN A  
; APPLICANT: LAZ, THOMAS M  
; APPLICANT: NAGORNY, RAISA  
; APPLICANT: WILSON, AMY E  
; TITLE OF INVENTION: DNA ENCODING A HUMAN MELANIN CONCENTRATING HORMONE RECEPTOR (MCH1  
; FILE REFERENCE: 1795/57453-A-PCT-US

; CURRENT APPLICATION NUMBER: US/09/885,478  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: PCT/US99/31169  
; PRIOR FILING DATE: 1999-12-30  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: ARTIFICIAL SEQUENCE  
; FEATURE:  
; OTHER INFORMATION: MUTATED MCH RECEPTOR  
US-09-885-478-28

Query Match 78.6%; Score 33; DB 4; Length 353;  
Best Local Similarity 75.0%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDRFLATV 9  
Db 139 IDRYLATV 146  
:|:|:|

## RESULT 15

US-09-919-039-127  
; Sequence 127, Application US/09919039  
; Patent No. 6727066  
; GENERAL INFORMATION:  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
; FILE REFERENCE: PA-0035 US  
; CURRENT APPLICATION NUMBER: US/09/919,039  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 60/222,113  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 401  
; SOFTWARE: PERL Program  
; SEQ ID NO 127  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6727066 1459432CD1  
US-09-919-039-127

Query Match 78.6%; Score 33; DB 4; Length 353;  
Best Local Similarity 75.0%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDRFLATV 9  
Db 139 IDRYLATV 146  
:|:|:|

Search completed: November 18, 2005, 09:36:01  
Job time : 15.6452 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 01:08:37 ; Search time 44.3226 Seconds  
(without alignments)  
84.961 Million cell updates/sec

Title: US-10-006-177-8

Perfect score: 42

Sequence: 1 LLDRFLATV 9

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	10	US-09-865-548A-151
2	42	100.0	9	14	US-10-006-177-8
3	42	100.0	9	17	US-10-705-459-151
4	42	100.0	10	17	US-10-705-459-321
5	42	100.0	254	9	US-09-796-149-4
6	42	100.0	377	11	US-09-736-250-1
7	42	100.0	377	17	US-10-485-555-30
8	36	85.7	139	15	US-10-424-599-277337
9	35	83.3	126	16	US-10-425-115-161998
10	35	83.3	645	16	US-10-739-930-10518
11	33	78.6	123	16	US-10-437-963-166119

12	33	78.6	190	17	US-10-732-923-2565	Sequence 2565, Ap
13	33	78.6	228	16	US-10-425-115-323325	Sequence 323325,
14	33	78.6	239	9	US-09-864-761-38414	Sequence 38414, A
15	33	78.6	256	14	US-10-156-761-13846	Sequence 13846, A
16	33	78.6	291	17	US-10-732-923-2912	Sequence 2912, Ap
17	33	78.6	345	14	US-10-309-515-52	Sequence 52, Appl
18	33	78.6	345	14	US-10-126-764-52	Sequence 52, Appl
19	33	78.6	347	14	US-10-309-515-60	Sequence 60, Appl
20	33	78.6	352	14	US-10-309-515-48	Sequence 48, Appl
21	33	78.6	352	14	US-10-126-764-48	Sequence 48, Appl
22	33	78.6	353	9	US-09-925-776-2	Sequence 2, Appli
23	33	78.6	353	9	US-09-895-686-2	Sequence 2, Appli
24	33	78.6	353	9	US-09-885-478-4	Sequence 4, Appli
25	33	78.6	353	9	US-09-885-478-28	Sequence 28, Appli
26	33	78.6	353	10	US-09-899-732-4	Sequence 4, Appli
27	33	78.6	353	10	US-09-964-923A-9	Sequence 9, Appli
28	33	78.6	353	10	US-09-919-039-127	Sequence 127, App
29	33	78.6	353	10	US-09-826-509-351	Sequence 351, App
30	33	78.6	353	10	US-09-826-509-355	Sequence 355, App
31	33	78.6	353	10	US-09-826-509-363	Sequence 363, App
32	33	78.6	353	10	US-09-826-509-367	Sequence 367, App
33	33	78.6	353	10	US-09-826-509-371	Sequence 371, App
34	33	78.6	353	10	US-09-826-509-375	Sequence 375, App
35	33	78.6	353	10	US-09-826-509-379	Sequence 379, App
36	33	78.6	353	10	US-09-826-509-383	Sequence 383, App
37	33	78.6	353	10	US-09-826-509-387	Sequence 387, App
38	33	78.6	353	14	US-10-029-314-4	Sequence 4, Appli
39	33	78.6	353	14	US-10-029-314-28	Sequence 28, Appli
40	33	78.6	353	14	US-10-221-461-2	Sequence 2, Appli
41	33	78.6	353	14	US-10-221-461-3	Sequence 3, Appli
42	33	78.6	353	14	US-10-221-461-4	Sequence 4, Appli
43	33	78.6	353	14	US-10-182-509-4	Sequence 4, Appli
44	33	78.6	353	14	US-10-309-515-2	Sequence 2, Appli
45	33	78.6	353	14	US-10-309-515-6	Sequence 6, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-865-548A-151  
; Sequence 151, Application US/09865548A  
; Publication No. US20030096298A1  
; GENERAL INFORMATION:  
; APPLICANT: Barnea, Eilon  
; APPLICANT: Beer, Ilan  
; APPLICANT: Ziv, Tamar  
; APPLICANT: Admon, Arie  
; TITLE OF INVENTION: METHOD OF IDENTIFYING PEPTIDES CAPABLE OF BINDING TO MHC MOLECULE  
; TITLE OF INVENTION: PEPTIDES IDENTIFIED THEREBY AND THEIR USES  
; FILE REFERENCE: 01/22080  
; CURRENT APPLICATION NUMBER: US/09/865.548A  
; CURRENT FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: US 60/290,958  
; PRIOR FILING DATE: 2001-05-16  
; NUMBER OF SEQ ID NOS: 204  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 151  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-09-865-548A-151

Query Match 100.0%; Score 42; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0

Qy 1 LLDRFLATV 9  
Db 1 LLDRFLATV 9

```
RESULT 2
US-10-006-177-8
; Sequence 8, Application US/10006177
; Publication No. US20030165513A1
; GENERAL INFORMATION:
; APPLICANT: Ramakrishna, Venky
; APPLICANT: Ross, Mark
; APPLICANT: Philip, Ramila
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treatm
; TITLE OF INVENTION: Diagnosis of Cancer
; FILE REFERENCE: 26747-35
; CURRENT APPLICATION NUMBER: US/10/006.177
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US/60/251,022
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US/60/256,824
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Epitopic Peptide
US-10-006-177-8

Query Match      100.0%; Score 42; DB 14; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9
Db 1 LLDRFLATV 9

RESULT 3
US-10-705-459-151
; Sequence 151, Application US/10705459
; Publication No. US20050053918A1
; GENERAL INFORMATION:
; APPLICANT: Barnea, Eilon
; APPLICANT: Beer, Ilan
; APPLICANT: Ziv, Tamar
; APPLICANT: Admon, Arie
; APPLICANT: Dassau, Lior
; APPLICANT: Buchsbaum, Samuel
; TITLE OF INVENTION: METHOD OF IDENTIFYING PEPTIDES CAPABLE OF BINDING TO MHC MOLECULE
; TITLE OF INVENTION: PEPTIDES IDENTIFIED THEREBY AND THEIR USES
; FILE REFERENCE: 26884
; CURRENT APPLICATION NUMBER: US/10/705.459
; CURRENT FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 372
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 151
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-705-459-151

Query Match      100.0%; Score 42; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9
Db 1 LLDRFLATV 9

RESULT 4
US-10-705-459-321
; Sequence 321, Application US/10705459
; Publication No. US20050053918A1
; GENERAL INFORMATION:
; APPLICANT: Barnea, Eilon
; APPLICANT: Beer, Ilan
; APPLICANT: Ziv, Tamar
; APPLICANT: Admon, Arie
; APPLICANT: Dassau, Lior
; APPLICANT: Buchsbaum, Samuel
; TITLE OF INVENTION: METHOD OF IDENTIFYING PEPTIDES CAPABLE OF BINDING TO MHC MOLECULE
; TITLE OF INVENTION: PEPTIDES IDENTIFIED THEREBY AND THEIR USES
; FILE REFERENCE: 26884
; CURRENT APPLICATION NUMBER: US/10/705.459
; CURRENT FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 372
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 321
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-705-459-321

Query Match      100.0%; Score 42; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9
Db 2 LLDRFLATV 10

RESULT 5
US-09-796-149-4
; Sequence 4, Application US/09796149
; Patent No. US20020035079A1
; GENERAL INFORMATION:
; APPLICANT: Univ. of Southern California
; TITLE OF INVENTION: Mutated cyclin G1 protein
; FILE REFERENCE: 4-31342A/USC
; CURRENT APPLICATION NUMBER: US/09/796.149
; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-149-4

Query Match      100.0%; Score 42; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9
Db 33 LLDRFLATV 41

RESULT 6
US-09-736-250-1
; Sequence 1, Application US/09736250
; Publication No. US20050014139A1
; GENERAL INFORMATION:
; APPLICANT: SUMITOMO ELECTRIC INDUSTRIES, LTD.
; APPLICANT: NAKAMURA, Takeshi
; TITLE OF INVENTION: HUMAN CYCLIN I AND GENES ENCODING THE SAME
; FILE REFERENCE: 050212-0278
; CURRENT APPLICATION NUMBER: US/09/736.250
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: 09/054,492
```



; PRIOR FILING DATE: 1998-04-03  
; PRIOR APPLICATION NUMBER: PCT/JP96/02905  
; PRIOR FILING DATE: 1996-10-07  
; PRIOR APPLICATION NUMBER: 284663/1995  
; PRIOR FILING DATE: 1995-10-05  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-736-250-1

Query Match 100.0%; Score 42; DB 11; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9  
| | | | |  
Db 72 LLDRFLATV 80

## RESULT 7

US-10-485-555-30  
; Sequence 30, Application US/10485555  
; Publication No. US20050120393A1  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies, Inc.  
; TITLE OF INVENTION: Full Length Prostate Specific Polynucleotides and Polypeptides  
; FILE REFERENCE: 9U 101.2 PCT  
; CURRENT APPLICATION NUMBER: US/10/485,555  
; CURRENT FILING DATE: 2004-02-03  
; PRIOR APPLICATION NUMBER: US 60/309,470  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: US 60/330,747  
; PRIOR FILING DATE: 2001-10-30  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-485-555-30

Query Match 100.0%; Score 42; DB 17; Length 377;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9  
| | | | |  
Db 72 LLDRFLATV 80

## RESULT 8

US-10-424-599-277337  
; Sequence 277337, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 277337  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_92459C.1.pap  
US-10-424-599-277337

Query Match 85.7%; Score 36; DB 15; Length 139;  
Best Local Similarity 77.8%; Pred. No. 27;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLATV 9  
| | | | |  
Db 81 LLDRFLATV 89

## RESULT 9

US-10-425-115-361998  
; Sequence 361998, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 361998  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(126)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_93318C.1.pap  
US-10-425-115-361998

Query Match 83.3%; Score 35; DB 16; Length 126;  
Best Local Similarity 87.5%; Pred. No. 38;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLAT 8  
| | | | |  
Db 97 VLDRFLAT 104

## RESULT 10

US-10-739-930-10518  
; Sequence 10518, Application US/10739930  
; Publication No. US20040216190A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; FILE REFERENCE: 38-21(53377)B  
; CURRENT APPLICATION NUMBER: US/10/739,930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO 10518  
; LENGTH: 645  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(645)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C460\_1.p  
US-10-739-930-10518

Query Match 83.3%; Score 35; DB 16; Length 645;

Best Local Similarity 87.5%; Pred. No. 2e+02; 0; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 0;

Qy 1 LLDRFLAT 8  
Db 300 LLDRFLST 307  
|||||:|  
|||||

## RESULT 11

US-10-437-963-166119  
; Sequence 166119, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbaruk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 166119  
; LENGTH: 123  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_6485C.1.pep  
US-10-437-963-166119

Query Match 78.6%; Score 33; DB 16; Length 123;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLA 7  
Db 46 LLDRFLA 52  
|||||  
|||||

## RESULT 12

US-10-732-923-2565  
; Sequence 2565, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 2565  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Stizostedion vitreum  
US-10-732-923-2565

Query Match 78.6%; Score 33; DB 17; Length 190;  
Best Local Similarity 87.5%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LDRFLATV 9  
Db 23 LDRFLAVV 30  
|||||  
|||||

## RESULT 13

US-10-425-115-323325  
; Sequence 323325, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 323325  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)-(228)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_57944C.1.pep  
US-10-425-115-323325

Query Match 78.6%; Score 33; DB 16; Length 228;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLDRFLA 7  
Db 177 LLDRFLA 183  
|||||  
|||||

## RESULT 14

US-09-864-761-38414  
; Sequence 38414, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 38414  
;; LENGTH: 239  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO Z86090.10  
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.67  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1  
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.78  
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.74  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3  
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.72  
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2  
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1  
;; OTHER INFORMATION: EST HUMAN HIT: BE701073.1, EVALUE 6.00e-44  
;; OTHER INFORMATION: SWISSPROT HIT: Q99705, EVALUE 0.00e+00  
US-09-864-761-38414

Query Match 78.6%; Score 33; DB 9; Length 239;  
Best Local Similarity 75.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LDRFLATV 9  
Db 25 IDRYLATV 32  
:::|||||

RESULT 15  
US-10-156-761-13846  
; Sequence 13846, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 13846  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-13846

Query Match 78.6%; Score 33; DB 14; Length 256;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LDRFLA 7  
Db 11 LDRFLA 17  
|||||

Search completed: November 18, 2005, 10:44:41  
Job time : 46.3226 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:03:06 ; Search time 59.6129 Seconds  
(without alignments)  
71.366 Million cell updates/sec

Title: US-10-006-177-9  
Perfect score: 53  
Sequence: 1 LLIDDKGTIKL 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	100.0	11	5	ABG32309 HLA-A2 as
2	53	100.0	270	2	AAY43932 Human pro
3	53	100.0	297	2	AAY95688 Homo sapi
4	53	100.0	297	2	AAY95687 Rattus no
5	53	100.0	297	4	AAG67613 Amino aci
6	53	100.0	297	4	AAG67434 Amino aci
7	53	100.0	297	6	ABU56522 Lung canc
8	53	100.0	297	6	ABR47405 Breast ca
9	53	100.0	297	6	AAE34491 Human CDC
10	53	100.0	297	7	ADE63825 Rat Prote
11	53	100.0	297	7	ADE63827 Human pro
12	53	100.0	297	7	ADF45037 Human kin
13	53	100.0	297	8	ADI27508 Human cel
14	53	100.0	297	8	ADJ66598 cdc2 prot
15	53	100.0	297	8	ADQ19629 Human sof
16	53	100.0	297	8	ADQ09234 Human CDC
17	53	100.0	297	8	ADR99149 Cell divi
18	53	100.0	337	8	ADN03938 Antipori
19	43	81.1	275	8	ADN20489 Bacterial
20	41	77.4	653	2	AAW38178 Arabidops
21	41	77.4	653	3	AAB28569 Arabidops
22	40	75.5	314	8	ADL04435 M. catarr
23	40	75.5	2197	8	ADK16573 Nanoarcha
24	39	73.6	103	3	AAG22581 Zea mays
25	39	73.6	222	4	AAY72763 Drosophil

26	39	73.6	277	6	ABU63704
27	39	73.6	277	6	ABU63703
28	39	73.6	277	7	ADD15786
29	39	73.6	277	7	ADD15787
30	39	73.6	430	7	ABO73870
31	39	73.6	492	4	ABB58135
32	39	73.6	537	5	ABB48708
33	39	73.6	661	3	AAI70098
34	39	73.6	690	2	AAAR40974
35	39	73.6	690	3	AAI70101
36	38	71.7	136	5	ABP43187
37	38	71.7	218	3	AAAB43679
38	38	71.7	225	3	AAAB51908
39	38	71.7	322	4	ABG26317
40	38	71.7	323	4	AAU30333
41	38	71.7	461	8	ADN21361
42	38	71.7	533	3	AAG27700
43	38	71.7	536	3	AAG27699
44	38	71.7	622	3	AAG27698
45	38	71.7	718	4	AAU35308

## ALIGNMENTS

RESULT 1  
ABG32309  
ID ABG32309 standard; peptide; 11 AA.  
XX AC ABG32309;  
XX DT 05-NOV-2002 (first entry)  
XX DE HLA-A2 associated immunogenic peptide from human CDC2 protein.  
XX KW Human; immunogen; epitope; HLA-A1; human leukocyte antigen; CTL;  
KW cytotoxic Y lymphocyte; cytostatic; cancer; colorectal carcinoma;  
KW ovarian carcinoma; lung carcinoma; prostate carcinoma; vaccine; tumour;  
XX HLA-2; passive immunotherapy; CDC2; cell division cycle 2.  
XX OS Homo sapiens.  
XX PN W0200246416-A2.  
XX PD 13-JUN-2002.  
XX PF 04-DEC-2001; 2001WO-US047290.  
XX PR 04-DEC-2000; 2000US-0251022P.  
XX PA 20-DEC-2000; 2000US-0256824P.  
XX PI (ARGO-) ARGONEX INC.  
XX PL Ramakrishna V, Ross M, Philip R;  
XX WPI; 2002-619021/66.  
XX DR New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte,  
XX and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.  
XX PT Claim 1; Page 50; 60pp; English.  
XX CC The invention relates to an immunogen comprising an isolated polypeptide  
CC whose amino acid sequence comprises an epitopic peptide, does not include  
CC MAGE 4 or MFG-E8 proteins, or consists of MAGE D protein or its  
CC immunologically active fragment. Also included are a polynucleotide  
CC encoding the immunogen or its complement, a vector comprising the  
CC polynucleotide, a mammalian cell comprising the vector and expressing the  
CC antibody specific for the immunogen. The immunogen is useful for inducing  
CC a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour  
CC cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is  
CC useful for inducing a CTL response when administered to a subject. A

CC mammalian cell that can express the immunogen, is useful for inducing a  
 CC CTL response in vitro that is specific for a tumour cell expressing HLA-1  
 CC or HLA-2. The immunogen or cell is useful for inducing CTL for treating a  
 CC subject with cancer (carcinoma, preferably colorectal carcinoma, ovarian  
 CC carcinoma, lung carcinoma and prostate carcinoma). The immunogen is also  
 CC useful for screening and diagnostic agents, for gene screening in  
 CC patients afflicted with cancer, for screening a sample for the presence  
 CC of CRLs that specifically recognise the corresponding epitopes, as a  
 CC diagnostic tool to evaluate the efficacy of the immunotherapeutic  
 CC treatments, to prepare class I MHC (major histocompatibility class)  
 CC tetramers which are utilised in conjunction with flow cytometry to  
 CC quantitate the frequency of peptide-specific CTL that are present in a  
 CC samples of lymphocytes from an individual, and for stimulating the  
 CC production of antibodies for use in passive immunotherapy, for use as  
 CC diagnostic reagents, and for use as reagents in other processes such as  
 CC affinity chromatography. The present sequence is an immunogenic epitope  
 CC of the invention derived from human CDC2 (cell division cycle 2) protein  
 CC  
 CC  
 CC

XX Sequence 11 AA;

Query Match 100.0%; Score 53; DB 5; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.002;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
 |||||  
 Db 1 LLIDDKGTIKL 11

RESULT 2  
 AAY43932  
 ID AAY43932 standard; protein; 270 AA.

XX AC AAY43932;

DT 21-DEC-1999 (first entry)

XX DE Human protein kinase #4.

XX KW Prediction; secondary structure; alignment; evolutionary conservation;  
 KW homology; periodicity; co-variation analysis; antigenic site;  
 KW site directed mutagenesis; interaction.

XX OS Homo sapiens.

XX FN US5958784-A.

XX PD 28-SEP-1999.

XX PF 25-MAR-1992; 92US-00857224.

XX PR 25-MAR-1992; 92US-00857224.

XX PA (BENN/) BENNER S A.

XX PI Benner SA;

XX DR WPI; 1999-570766/48.

XX PT Predicting the folded structure of proteins.

XX PS Disclosure; Col 199-202; 113pp; English.

XX Sequences AAY43902-Y44015 represent proteins used in a novel method of  
 CC predicting the folded structure of proteins, by aligning sequences of  
 CC homologous proteins and using patterns of evolutionarily conserved and  
 CC varied sequences to assign positions. Positions in the alignment are  
 CC assigned to the surface or inside of the folded structure, active sites,  
 CC and parsing segments. Secondary structural units are assigned by  
 CC identifying periodicity in the assignments, and assembled into globular  
 CC form using distance constraints imposed by disulfide bridges, active site  
 CC assignments and co-variation analysis. The predicted secondary structures  
 CC are useful for identifying antigenic sites on a protein molecule, as

CC guides for site directed mutagenesis studies, and for understanding the  
 CC interaction of a protein with other molecules  
 CC  
 CC Sequence 270 AA;

Query Match 100.0%; Score 53; DB 2; Length 270;  
 Best Local Similarity 100.0%; Pred. No. 0.075;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
 |||||  
 Db 133 LLIDDKGTIKL 143

RESULT 3  
 AAW95688  
 ID AAW95688 standard; protein; 297 AA.

XX AC AAW95688;

XX DT 08-JUN-1999 (first entry)

XX DE Homo sapiens Cdc2 protein kinase.

XX KW cdc2; protein kinase; diagnosis; infection; immunosuppression; AIDS;  
 KW acquired immunodeficiency syndrome; diagnosis; therapy;  
 KW organ transplantation.

XX OS Homo sapiens.

XX PN WO9856799-A1.

XX PD 17-DEC-1998.

XX PF 12-JUN-1998; 98WO-US012100.

XX PR 13-JUN-1997; 97US-00874347.

XX PA (MAYO-) MAYO FOUNDATION.

XX PI Limper AH, Leof EB, Thomas CF, Gustafson MP;

XX DR WPI; 1999-080877/07.

XX PT New isolated polynucleotide encoding a Pneumocystis carinii Cdc2  
 FT polypeptide - useful for diagnosing or monitoring P. carinii infection in  
 FT patients with chronic immunosuppression.

XX PS Example; Page 42-43; 67pp; English.

XX The sequence is that of a Cdc2 polypeptide. The Cdc2 polypeptides are  
 CC useful in treating P. carinii pneumonia in patients with chronic  
 CC immunosuppression. Conditions associated with P. carinii include acquired  
 CC immune deficiency syndrome (AIDS), solid tumors, hematological  
 CC malignancies, organ transplantation, and inflammatory conditions. The  
 CC polypeptides are additionally useful in identifying agents that inhibit  
 CC the phosphorylation activity of P. carinii Cdc2 polypeptides. The  
 CC polypeptide and its antibodies are useful for diagnosing or monitoring P.  
 CC carinii infection in patients. The P. carinii Cdc2 polypeptides allow a  
 CC therapeutic approach to treating P. carinii pneumonia because they are  
 CC not limited by significant side effects. Polynucleotides encoding  
 CC P. carinii Cdc2 enable in vitro production of the protein, which can be  
 CC used in studying the organism (especially its life cycle) and developing  
 CC new therapies

XX Sequence 297 AA;

Query Match 100.0%; Score 53; DB 2; Length 297;  
 Best Local Similarity 100.0%; Pred. No. 0.083;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
 |||||

```

Db      134 LLIDDKGTIKL 144

RESULT 4
AAW95687
ID      AAW95687 standard; protein; 297 AA.
XX
AC      AAW95687;
XX
DT      08-JUN-1999 (first entry)
XX
DE      Rattus norvegicus Cdc2 protein kinase.
XX
KW      cdc2; protein kinase; diagnosis; infection; immunosuppression; AIDS;
KW      acquired immunodeficiency syndrome; diagnosis; therapy;
KW      organ transplantation.
XX
OS      Rattus norvegicus.
XX
PN      WO9856799-A1.
XX
PD      17-DEC-1998.
XX
PF      12-JUN-1998; 98WO-US012100.
XX
PR      13-JUN-1997; 97US-00874347.
XX
PA      (MAYO-) MAYO FOUNDATION.
XX
PI      Limper AH, Leof EB, Thomas CF, Gustafson MP;
XX
PS      WPI; 1999-080877/07.
XX
PT      New isolated polynucleotide encoding a Pneumocystis carinii Cdc2
PT      polypeptide - useful for diagnosing or monitoring P. carinii infection in
PT      patients with chronic immunosuppression.
XX
PS      Example; Page 42; 67pp; English.
XX
CC      The sequence is that of a Cdc2 polypeptide. The Cdc2 polypeptides are
CC      useful in treating P. carinii pneumonia in patients with chronic
CC      immunosuppression. Conditions associated with P. carinii include acquired
CC      immune deficiency syndrome (AIDS), solid tumors, hematological
CC      malignancies, organ transplantation, and inflammatory conditions. The
CC      polypeptides are additionally useful in identifying agents that inhibit
CC      the phosphorylation activity of P. carinii Cdc2 polypeptides. The
CC      polypeptide and its antibodies are useful for diagnosing or monitoring P.
CC      carinii infection in patients. The P. carinii Cdc2 polypeptides allow a
CC      therapeutic approach to treating P. carinii pneumonia because they are
CC      not limited by significant side effects. Polynucleotides encoding
CC      P. carinii Cdc2 enable in vitro production of the protein, which can be
CC      used in studying the organism (especially its life cycle) and developing
CC      new therapies
XX
SQ      Sequence 297 AA;

Query Match      100.0%; Score 53; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
        |||||
Db      134 LLIDDKGTIKL 144

RESULT 5
AAG67613
ID      AAG67613 standard; protein; 297 AA.
XX
AC      AAG67613;
XX
DT      26-NOV-2001 (first entry)
XX
DE      Amino acid sequence of a human polypeptide.
XX
KW      Human; protein kinase; protein phosphatase; signal transduction;
KW      intracellular signalling pathway.
XX
OS      Homo sapiens.
XX
PN      WO200109345-A1.
XX
PD      08-FEB-2001.
XX
XX
DE      Amino acid sequence of a human protein.
XX
KW      Human; protein kinase; protein phosphatase; signal transduction.
XX
OS      Homo sapiens.
XX
PN      WO200109316-A1.
XX
PD      08-FEB-2001.
XX
PF      28-JUL-2000; 2000WO-JP005061.
XX
PR      29-JUL-1999; 99JP-00248036.
PR      18-OCT-1999; 99US-0159590P.
PR      11-JAN-2000; 2000JP-00118776.
PR      17-FEB-2000; 2000US-0183322P.
PR      02-MAY-2000; 2000JP-00183767.
PR      09-JUN-2000; 2000JP-00241899.
XX
PA      (HELI-) HELIX RES INST.
XX
PI      Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI      Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otauki T, Funahashi S;
PI      Senoo C, Nezu J;
XX
PS      WPI; 2001-570286/64.
XX
PT      New genes encoding proteins with protein kinase/protein phosphatase
PT      activity, useful in the diagnosis and treatment of diseases.
XX
PS      Example 4; Page 91-94; 233pp; Japanese.
XX
CC      The specification describes human protein kinase/protein phosphatases. It
CC      is expected that the protein kinase/protein phosphatase gene participates
CC      in signal transduction in cells. The protein kinase/protein phosphatase
CC      polypeptides and polynucleotides are useful for developing diagnostics
CC      and treatment agents for human and animal diseases. The protein
CC      kinase/protein phosphatase polypeptides are useful as target molecules in
CC      designing novel drugs. The protein kinase/protein phosphatase
CC      polynucleotides are useful as a source of probes and primers, which may
CC      be used to isolate homologous sequences. The present sequence represents
CC      a human protein, which is used in the course of the invention
XX
SQ      Sequence 297 AA;

Query Match      100.0%; Score 53; DB 4; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
        |||||
Db      134 LLIDDKGTIKL 144

RESULT 6
AAG67434
ID      AAG67434 standard; protein; 297 AA.
XX
AC      AAG67434;
XX
DT      26-NOV-2001 (first entry)
XX
DE      Amino acid sequence of a human polypeptide.
XX
KW      Human; protein kinase; protein phosphatase; signal transduction;
KW      intracellular signalling pathway.
XX
OS      Homo sapiens.
XX
PN      WO200109345-A1.
XX
PD      08-FEB-2001.
XX
XX

```

```
PF 28-JUL-2000; 2000WO-JP005060.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 18-OCT-1999; 99US-0159590P.
PR 11-JAN-2000; 2000JP-00118776.
PR 17-FEB-2000; 2000US-0183332P.
PR 02-MAY-2000; 2000JP-00183767.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
PI Senoo C, Nezu J;
XX
XX WPI; 2001-564736/63.
XX
XX New genes encoding protein kinase and protein phosphatase, useful for
PT identifying modulators which can be used to treat human or animal
PT disorders associated with the expression or function of these enzymes.
XX
XX Example 4; Page 193-195; 336pp; Japanese.
XX
XX The specification describes human protein kinase/protein phosphatases.
CC The polypeptides are expected to participate in signal transduction in
CC cells. The kinase phosphatases are connected with intracellular
CC signalling pathways. Antisense oligonucleotides and compounds identified
CC by screening (agonists or antagonists) can be used to treat human or
CC animal disorders associated with the expression or function of the
CC protein. In addition, the polypeptides may be used as target molecules
CC for drug development. The present sequence represents a polypeptide, used
CC in the course of the invention
XX
XX Sequence 297 AA;
SQ
Query Match 100.0%; Score 53; DB 4; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144
RESULT 7
ABUS6522
ID ABUS6522 standard; protein; 297 AA.
XX
XX ABUS6522;
AC
XX
DT 02-APR-2003 (first entry)
XX
XX Lung cancer-associated polypeptide #115.
XX
XX Lung cancer-associated polypeptide; cytostatic; emphysema;
XX antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
XX Unidentified.
XX
XX WO200286443-A2.
XX
XX 31-OCT-2002.
XX
XX 18-APR-2002; 2002WO-US012476.
XX
XX 18-APR-2001; 2001US-0284770P.
XX 10-MAY-2001; 2001US-0290492P.
XX 09-NOV-2001; 2001US-0339245P.
XX 13-NOV-2001; 2001US-0350666P.
XX 29-NOV-2001; 2001US-0334370P.
XX 12-APR-2002; 2002US-0372246P.
XX
(EOB-) EOS BIOTECHNOLOGY INC.
XX
XX Aziz N, Murray R;
XX
XX WPI; 2003-093161/08.
XX N-PSDB; ABX76248.
XX
XX Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
XX Claim 27; Page 280-281; 453pp; English.
XX
XX The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
CC invention
XX
XX Sequence 297 AA;
SQ
Query Match 100.0%; Score 53; DB 6; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144
RESULT 8
ABR47405
ID ABR47405 standard; protein; 297 AA.
XX
XX ABR47405;
AC
XX
XX 12-JUN-2003 (first entry)
XX
XX Breast cancer associated protein sequence SEQ ID NO:41.
XX
XX Human; breast cancer; cytostatic; gene therapy.
XX
XX Homo sapiens.
XX
XX WO2003004989-A2.
XX
XX 16-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-US019669.
XX
XX 21-JUN-2001; 2001US-0299887P.
XX 27-JUN-2001; 2001US-0301572P.
XX 18-JUL-2001; 2001US-0306501P.
XX 25-SEP-2001; 2001US-0325002P.
XX 05-MAR-2002; 2002US-0362585P.
XX 14-MAY-2002; 2002US-0380391P.
XX
```



PA (MILL-) MILLENIUM PHARM INC.  
 XX  
 PI Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;  
 PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;  
 PI Bast RC, Hortobagyi GN, Fuszai L, Mexic F, Sahin A, Mills GB;  
 XX  
 DR WPI; 2003-210381/20.  
 DR N-PSDB; ACC50096.  
 XX  
 XX Breast cancer diagnosis or treatment by comparing the level of expression  
 PT of a marker in a patient sample with that in the control non-breast  
 PT cancer sample.  
 XX  
 PS Claim 1; SEQ ID NO 41; 128pp; English.  
 XX  
 CC The present invention describes a method for assessing whether a patient  
 CC is afflicted with breast cancer. The method comprises comparing the level  
 CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and  
 CC ABR47396 to ABR47632) in a patient sample and the normal level of  
 CC expression of the marker in a control non-breast cancer sample, where a  
 CC significant increase in the level of expression of the marker in the  
 CC patient sample and the normal level is an indication that the patient is  
 CC afflicted with breast cancer. The breast cancer associated sequences from  
 CC the present invention have cytostatic activities and can be used in gene  
 CC therapy. The method is useful for diagnosing and treating breast cancer.  
 CC N.B. The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 297 AA;  
 Query Match 100.0%; Score 53; DB 6; Length 297;  
 Best Local Similarity 100.0%; Pred. No. 0.083;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 LLIDDKGITKL 11  
 Db 134 LLIDDKGITKL 144  
 RESULT 9  
 AAEE34491  
 ID AAEE34491 standard; protein; 297 AA.  
 XX  
 AC AAEE34491;  
 XX  
 XX 14-MAY-2003 (first entry)  
 XX  
 XX Human CDC2 protein.  
 XX  
 KW Human; pyridylpyrimidine derivative; cellular protein kinase; Scrapie;  
 KW cellular protein phosphatase; cellular signal transduction; prophylaxis;  
 KW prion infection; chronic wasting disease; CWD; Creutzfeldt-Jacob disease;  
 KW CJD; transmissible mink encephalopathy; bovine spongiform encephalopathy;  
 KW TME; BSE; Gerstmann-Strausler-Scheinker syndrome; GSS; Alpers syndrome;  
 KW fatal familial insomnia; FFI; kuru; neurodegenerative disease; nontropic;  
 KW Alzheimer's disease; CDC-2; CDK1.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200293164-A2.  
 XX  
 XX 21-NOV-2002.  
 XX  
 XX 16-MAY-2002; 2002WO-EP005420.  
 XX  
 XX 16-MAY-2001; 2001EP-00111858.  
 PR 29-MAY-2001; 2001US-0293528P.  
 PR 13-JUL-2001; 2001EP-00117113.  
 PR 18-JUL-2001; 2001US-0305898P.  
 XX  
 XX (AXXI-) AXXINA PHARM AG.  
 PA  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 PI Stein-Gerlach M, Salassidis K, Bacher G, Mueller S;  
 XX  
 DR WPI; 2003-120714/11.  
 DR N-PSDB; AAD52783.  
 XX  
 PT New pyridylpyrimidine derivatives useful in the treatment or prevention  
 PT of infectious disease e.g. Kuru syndrome and Creutzfeldt-Jacob disease  
 PT (CJD).  
 XX  
 PS Disclosure; Page 75-76; 96pp; English.  
 XX  
 CC The invention relates to novel pyridylpyrimidine derivatives and methods  
 CC of detecting prion infections and/or prion disease in an individual or in  
 CC cells, cell cultures and/or cell lysates. The method involves adding at  
 CC least one monoclonal or polyclonal antibody, oligonucleotide or pyridyl-  
 CC pyrimidine derivative to the sample or in cells, cell cultures and/or  
 CC cell lysates and detecting the activity of at least one human cellular  
 CC protein kinases (e.g., GGF-R1 (also known as fig, Fl-1, Flt-2, b-FGFR),  
 CC Tkt (also known as CCK-2, DDR-2 or EDDR; EC number 2.7.1.112), Abl (also  
 CC known as c-abl), ctk1, MKK7 (also known as SAPK1a, SAPKalpha), CDC2 (also  
 CC known as CDK1), PRK), human cellular protein phosphatases such as PTP-SL  
 CC (also known as MCP83) and PTP-zeta, the cellular signal transduction  
 CC molecules HSP80 and GPCR-1. The invention is useful for regulating the  
 CC production of prions in cells and in the manufacture of pharmaceutical  
 CC composition for prophylaxis and/or treatment of infectious disease (e.g.  
 CC scrapie, chronic wasting disease (CWD), transmissible mink encephalopathy  
 CC (TME), Creutzfeldt-Jacob disease (CJD), bovine spongiform encephalopathy  
 CC (BSE), variant CJD, Gerstmann-Strausler-Scheinker syndrome (GSS), fatal  
 CC familial insomnia (FFI), Kuru and Alpers syndrome, especially BSE, CJD  
 CC vCJD) or neurodegenerative diseases (e.g., Alzheimer's disease) in humans  
 CC or ruminants. The present sequence is human CDC2 protein used in the  
 CC invention  
 XX  
 SQ Sequence 297 AA;  
 Query Match 100.0%; Score 53; DB 6; Length 297;  
 Best Local Similarity 100.0%; Pred. No. 0.083;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 LLIDDKGITKL 11  
 Db 134 LLIDDKGITKL 144  
 RESULT 10  
 ADE63825  
 ID ADE63825 standard; protein; 297 AA.  
 XX  
 XX ADE63825;  
 AC ADE63825;  
 XX  
 XX 29-JAN-2004 (first entry)  
 XX  
 XX Rat Protein P39951, SEQ ID NO 9769.  
 DE  
 XX  
 XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
 KW  
 XX Rattus norvegicus.  
 OS  
 XX WO2003016475-A2.  
 XX  
 XX 27-FEB-2003.  
 XX  
 XX 14-AUG-2002; 2002WO-US025765.  
 XX  
 XX 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 PA  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;

```

XX WPI; 2003-268312/26.
DR GENBANK; P39951.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 297 AA;
SQ
Query Match 100.0%; Score 53; DB 7; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144
RESULT 11
ADE63827
ID ADE63827 standard; protein; 297 AA.
XX
XX ADE63827;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human Protein P06493, SEQ ID NO 9771.
DE
XX
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
OS
XX
XX WO2003016475-A2.
FN
XX
XX 27-FEB-2003.
PD
XX
XX 14-AUG-2002; 2002WO-US025765.
PF
XX
XX 14-AUG-2001; 2001US-0312147P.
PR
XX
XX 01-NOV-2001; 2001US-0346382P.
PR
XX
XX 26-NOV-2001; 2001US-0333347P.
PR

```

---

```

XX (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX WPI; 2003-268312/26.
DR GENBANK; P06493.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 297 AA;
SQ
Query Match 100.0%; Score 53; DB 7; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144
RESULT 12
ADF45037
ID ADF45037 standard; protein; 297 AA.
XX
XX ADF45037;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX
XX Human kinase CDK1.
DE
XX
XX Human; protein kinase; enzyme; inhibitor; CDK1.
KW
XX
XX Homo sapiens.
OS
XX
XX WO2003081210-A2.
FN
XX
XX 02-OCT-2003.
PD
XX
XX 20-MAR-2003; 2003WO-US008725.
PF
XX
XX

```

PR 21-MAR-2002; 2002US-0366892P.  
XX (SUNE-) SUNESIS PHARM INC.  
XX Prescott JC, Braisted A;  
XX WPI; 2003-865136/80.  
XX Identifying ligand binding to inactive conformation of target protein  
PT kinase (T) comprises contacting the conformation modified (T) which  
PT contains reactive group at binding site, with ligands and detecting  
PT kinase-ligand conjugate formation.  
XX Disclosure; SEQ ID NO 6; 260pp; English.  
XX The present invention relates to a method for identifying a ligand (L),  
CC which binds to an inactive conformation of target protein kinase (T). The  
CC method involves contacting inactive conformation of (T), which contains  
CC or is modified to contain a reactive group at or near a binding site of  
CC interest, with one or more ligand candidates capable of covalently  
CC bonding to the reactive group thus forming a kinase-(L) conjugate (C).  
CC The method is useful for identifying protein kinase inhibitors that  
CC preferentially bind to inactive conformation of a target protein kinase.  
CC The present sequence is a protein kinase which may be modified via an  
CC amino acid substitution, for use in the method of the invention.  
XX  
XX Sequence 297 AA;  
SQ  
Query Match 100.0%; Score 53; DB 7; Length 297;  
Best Local Similarity 100.0%; Pred. No. 0.083;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LLIDDKGTIKL 11  
Db 134 LLIDDKGTIKL 144  
RESULT 13  
AD127508  
ID AD127508 standard; protein; 297 AA.  
XX  
XX AD127508;  
XX  
XX 22-APR-2004 (first entry)  
XX Human cell division cycle 2.  
XX hyperproliferative disorder; cancer; bone metabolism;  
XX Alzheimer's disease; human; cell division cycle 2.  
XX Homo sapiens.  
XX US2004006029-A1.  
XX 08-JAN-2004.  
XX 02-JUL-2002; 2002US-00189266.  
XX 02-JUL-2002; 2002US-00189266.  
XX (ISIS-) ISIS PHARM INC.  
XX Dean NM, Freier SM;  
XX WPI; 2004-081741/08.  
XX N-ESDB; AD127361.  
XX New compounds, particularly antisense oligonucleotides targeted to a  
PT nucleic acid encoding cell division cycle 2, useful for treating cancer,  
PT a disease resulting from dysregulation of bone metabolism or Alzheimer's  
PT disease.  
XX Disclosure; Page 36-67; 61pp; English.

XX The invention relates to a compound targeted to and which specifically  
CC hybridizes with a nucleic acid molecule encoding cell division cycle 2,  
CC and inhibits the expression of cell division cycle 2. The compound,  
CC composition and methods are useful for treating a disease or condition  
CC associated with cell division cycle 2, such as hyperproliferative  
CC disorder e.g. cancer, or a disease or condition resulting from  
CC dysregulation of bone metabolism, or Alzheimer's disease. They are also  
CC useful in research and diagnostics for modulating the expression of cell  
CC division cycle 2. The present sequence represents the amino acid sequence  
CC of a human cell division cycle 2.  
XX  
XX Sequence 297 AA;  
SQ  
Query Match 100.0%; Score 53; DB 8; Length 297;  
Best Local Similarity 100.0%; Pred. No. 0.083;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LLIDDKGTIKL 11  
Db 134 LLIDDKGTIKL 144  
RESULT 14  
ADJ66598  
ID ADJ66598 standard; protein; 297 AA.  
XX  
XX ADJ66598;  
XX  
XX 06-MAY-2004 (first entry)  
XX cdc2 protein for anti-cancer protein complex.  
XX neuroprotective; cytostatic; gene therapy; protein complex;  
XX cellular network; cancer; neurodegenerative disease; drug target.  
XX Homo sapiens.  
XX WO2004009622-A2.  
XX 29-JAN-2004.  
XX 18-JUL-2003; 2003WO-EP007835.  
XX 19-JUL-2002; 2002EP-00016109.  
XX 19-JUL-2002; 2002EP-00016111.  
XX 19-JUL-2002; 2002EP-00016123.  
XX 19-JUL-2002; 2002EP-00016128.  
XX 22-JUL-2002; 2002EP-00016427.  
XX (CELL-) CELLZONE AG.  
XX Merino A, Bouwmeester T, Bauer A, Drewes G, Marzioch M, Kruse U;  
XX Superti-Furga G, Eberhard D, Ruffner H, Hobson S, Helftenbein G;  
XX Cruciat C;  
XX WPI; 2004-123372/12.  
XX New protein complexes of cellular networks underlying the development of  
PT cancer and other diseases, useful for diagnosing and/or treating  
PT neurodegenerative diseases or cancer, and in drug screening.  
XX Disclosure; SEQ ID NO 128; 809pp; English.  
XX The invention relates to a protein complex of cellular networks  
CC underlying the development of cancer and other diseases. The complex (I)  
CC comprises at least one first and second proteins selected from any of the  
CC proteins listed in the specification, or their functionally active  
CC derivatives, fragments, homologues or variants, the variants being  
CC encoded by a nucleic acid that hybridizes to the nucleic acid encoding  
CC the protein under low stringency conditions. A complex (II) comprises at  
CC least two of the second proteins, where the low stringency conditions  
CC comprise hybridization in a buffer comprising 35% formamide, 5 x SSC, 50

CC mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% BSA, 100 microg/ml  
CC denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20  
CC hours at 40 deg C, washing in a buffer consisting of 2 x SSC, 25 mM Tris-  
CC HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 55 deg C, and  
CC washing in a buffer consisting of 2 x SSC, 25 mM Tris-HCl (pH 7.4), 5 mM  
CC EDTA, and 0.1% SDS for 1.5 hours at 60 deg C. The composition and methods  
CC are useful in diagnosing or treating diseases and disorders, preferably  
CC neurodegenerative diseases. These may also be used as a drug target or in  
CC manufacturing a medicament for the treatment or prevention of the above-  
CC mentioned diseases or disorders. The composition may also be used for  
CC treating cancer. This sequence represents one of the proteins of the  
CC complex of the invention.

XX SQ Sequence 297 AA;

Query Match 100.0%; Score 53; DB 8; Length 297;  
Best Local Similarity 100.0%; Pred. No. 0.083;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
Db 134 LLIDDKGTIKL 144  
|||||

RESULT 15  
ADQ19629  
ID ADQ19629 standard; protein; 297 AA.

XX AC ADQ19629;

XX DT 26-AUG-2004 (first entry)

XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 2448.

XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.

XX OS Homo sapiens.

XX PN WO2004048938-A2.

XX PD 10-JUN-2004.

XX PF 26-NOV-2003; 2003WO-US038193.

XX PR 26-NOV-2002; 2002US-0429739P.

XX PA (PROT-) PROTEIN DESIGN LABS INC.

XX PI Aziz N, Ginsburg WM, Zlotnik A;

XX DR WPI; 2004-441208/41.

XX PT Early detection of soft tissue sarcoma comprises determining expression  
XX of a gene in a first soft tissue sample and a normal soft tissue sample  
XX and comparing the gene expression, also useful in treating soft tissue  
XX sarcoma.

XX PS Example 2; SEQ ID NO 2448; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma  
XX which comprises obtaining a first soft tissue sample from an individual  
XX and a normal soft tissue sample from the same or different individual,  
XX determining the expression of a gene in both samples and comparing the  
XX expression of the gene in both soft tissue samples, where a higher level  
XX of protein expression in the first soft tissue sample indicates the  
XX presence of soft tissue sarcoma. The method of the invention has  
XX cytostatic applications and may be useful for detecting soft tissue  
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic  
XX acid sequences may be useful in diagnostic and screening applications.  
XX The current sequence is that of a human soft tissue sarcoma-upregulated  
XX protein of the invention. The current sequence is not shown within the  
XX specification per se but was submitted in CD format by the inventor.

SQ Sequence 297 AA;

Query Match 100.0%; Score 53; DB 8; Length 297;  
Best Local Similarity 100.0%; Pred. No. 0.083;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
Db 134 LLIDDKGTIKL 144  
|||||

Search completed: November 18, 2005, 01:01:25  
Job time : 64.6129 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:11:06 ; Search time 10.8817 Seconds  
(Without alignments)  
97.263 Million cell updates/sec

Title: US-10-006-177-9

Perfect score: 53

Sequence: 1 LLIDDKGTIKL 11

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	100.0	297	1 A36074	protein kinase (EC
2	53	100.0	297	1 S24913	protein kinase (EC
3	53	100.0	297	2 A29539	protein kinase (EC
4	53	100.0	297	2 I45977	cyclin-dependent k
5	48	90.6	303	1 S86011	protein kinase (EC
6	43	81.1	302	1 I50474	protein kinase (EC
7	43	81.1	302	2 B44349	protein kinase (EC
8	42	79.2	302	1 A44349	protein kinase (EC
9	42	79.2	502	2 T02306	probable protein k
10	41	77.4	69	2 T07875	protein kinase hom
11	41	77.4	653	2 T47581	SCARECROW - Arabi
12	41	77.4	653	2 T51244	SCARECROW protein
13	40	75.5	100	2 P82469	hypothetical prote
14	40	75.5	355	2 P02077	protein kinase (EC
15	40	75.5	385	2 S49752	homeotic protein Y
16	40	75.5	1025	2 C82886	hypothetical prote
17	39	73.6	537	2 AH1277	probable transport
18	39	73.6	537	2 AH1640	probable transport
19	39	73.6	652	2 H86221	hypothetical prote
20	39	73.6	706	2 A48084	STE11 protein kina
21	38	71.7	124	2 A97132	hypothetical prote
22	38	71.7	480	2 A96371	hypothetical prote
23	38	71.7	542	2 T00887	protein kinase hom
24	38	71.7	596	1 B44238	fumarate reductase
25	38	71.7	710	2 T13458	hypothetical prote
26	37	69.8	69	2 T07876	protein kinase hom
27	37	69.8	151	2 S17451	cdc2 protein homol
28	37	69.8	200	2 AC2173	iron superoxide di
29	37	69.8	233	1 B69513	conserved hypothet

## RESULT 1

A36074

protein kinase (EC 2.7.1.37) cdc2 [validated] - mouse

N:Alternate names: cell division control protein cdc2; CTD kinase; p34 protein kinase  
C:Species: Mus musculus (house mouse)

C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004

C:Accession: A36074; S05049; I48288

R:Th'ng, J.P.H.; Wright, P.S.; Hamaguchi, J.; Lee, M.G.; Norbury, C.J.; Nurse, P.; Stradl Cell 63, 313-324, 1990

A:Title: The F210 cell line is a mouse G2 phase mutant with a temperature-sensitive CDC

A:Reference number: A36074; MUID:91004239; PMID:2208288

A:Accession: A36074

A:Molecule type: mRNA

A:Residues: 1-297 <THA>

A:Cross-references: UNIPROT:P11440; GB:M38724

R:Cisek, L.J.; Cordien, J.L.

Nature 339, 679-684, 1989

A:Title: Phosphorylation of RNA polymerase by the murine homologue of the cell-cycle con

A:Reference number: S05049; MUID:89295557; PMID:2662013

A:Accession: S05049

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-103,'T',105-112,'M',114-211,'Q',213-244,'N',246-259,'C',261-262,'F',264-2

A:Note: part of this sequence was confirmed by protein sequencing

R:Spurr, N.K.; Gough, A.C.; Lee, M.G.

DNA Seq. 1, 49-54, 1990

A:Title: Cloning of the mouse homologue of the yeast cell cycle control gene cdc2.

A:Reference number: I48288; MUID:92119319; PMID:2132958

A:Accession: I48288

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-164,'L',166-211,'Q',213-272,'T',274-297 <RES>

A:Cross-references: EMBL:X16461; NID:950359; PIDN:CAA34481.1; PID:G50360

C:Comment: The carboxyl-terminal domain of the large chain of DNA-dependent RNA polymera

specific kinase.

C:Genetics:

A:Gene: CDC2

C:Complex: In various organisms, cdc2 has been identified as a component of the M-phase

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonin

F:2-256/Domain: protein kinase homology <KIN>

F:10-18/Region: protein kinase ATP-binding motif

F:33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 100.0%; Score 53; DB 1; Length 297;

Best Local Similarity 100.0%; Pred. No. 0.046;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LLIDDKGTIKL 11

Db 134 LLIDDKGTIKL 144

```
RESULT 2
S24913
protein kinase (EC 2.7.1.37) cdc2 [similarity] - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: S24913
R;Kanaoka, Y.; Nojima, H.; Okayama, H.
submitted to the EMBL Data Library, July 1991
A;Description: Nucleotide sequences of cDNAs encoding rat cdc2 + and cyclin 2.
A;Reference number: S20658
A;Accession: S24913
A;Molecule type: mRNA
A;Residues: 1-297 <KAN>
A;Cross-references: UNIPROT:P39951; EMBL:X60767; NID:G57533; PIDN:CAA43177.1; PID:G57534
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k
F;2-256/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 100.0%; Score 53; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144

RESULT 3
A29539
protein kinase (EC 2.7.1.37) cdc2 - human
N;Alternate names: cell division control protein 2 (CDC2)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C;Accession: A29539
R;Lee, M.G.; Nurse, P.
Nature 327, 31-35, 1987
A;Title: Complementation used to clone a human homologue of the fission yeast cell cycle
A;Reference number: A29539; MUID:87201915; PMID:3553962
A;Accession: A29539
A;Molecule type: mRNA
A;Residues: 1-297 <LEE>
A;Cross-references: UNIPROT:P06493; GB:X05360; NID:G29838; PIDN:CAA28963.1; PID:G29839
C;Genetics:
A;Gene: GDB:CDC2
A;Cross-references: GDB:119052; OMIM:116940
A;Map position: 10q21.1-10q21.1
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine
F;2-256/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;14,161/Binding site: phosphate (Thr) (covalent) #status predicted
F;15/Binding site: phosphate (Tyr) (covalent) #status predicted
F;33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 100.0%; Score 53; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144

RESULT 4
I45977
cyclin-dependent kinase 1 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
C;Accession: I45977
R;Yang, L.; Farin, C.E.
```

```
Gene 141, 283-286, 1994
A;Title: Identification of cDNAs encoding bovine cyclin B and Cdk1/Cdc2.
A;Reference number: I45977; MUID:94215918; PMID:8163203
A;Accession: I45977
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-297 <VAN>
A;Cross-references: UNIPROT:P48734; GB:L26547; NID:G433155; PIDN:AAA18894.1; PID:G498173
C;Genetics:
A;Gene: cdk1/cdc2
C;Superfamily: kinase-related transforming protein; protein kinase homology
F;2-256/Domain: protein kinase homology <KIN>

Query Match 100.0%; Score 53; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144

RESULT 5
S06011
protein kinase (EC 2.7.1.37) cdc2 - chicken
N;Alternate names: cell division control protein cdc2
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S06011
R;Krek, W.; Nigg, E.A.
EMBO J. 8, 3071-3078, 1989
A;Title: Structure and developmental expression of the chicken CDC2 kinase.
A;Reference number: S06011; MUID:90059893; PMID:2684635
A;Accession: S06011
A;Molecule type: mRNA
A;Residues: 1-303 <KRE>
A;Cross-references: UNIPROT:P13863; EMBL:X16881; NID:G63172; PIDN:CAA34764.1; PID:G63173
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k
F;2-256/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 90.6%; Score 48; DB 1; Length 303;
Best Local Similarity 90.9%; Pred. No. 0.39;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144

RESULT 6
I50474
protein kinase (EC 2.7.1.37) cdc2 [similarity] - goldfish
C;Species: Carassius auratus (goldfish)
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: I50474
R;Kajiwara, H.; Yamashita, M.; Katsu, Y.; Nagahama, Y.
Dev. Growth Differ. 35, 647-654, 1993
A;Title: Isolation and characterization of goldfish cdc2, a catalytic component of matur
A;Reference number: I50474
A;Accession: I50474
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-302 <KAJ>
A;Cross-references: UNIPROT:P51958; GB:D17758; NID:G471097; PIDN:BAA04605.1; PID:G471098
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;2-256/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted
```

Query Match 81.1%; Score 43; DB 1; Length 302;  
 Best Local Similarity 81.8%; Pred. No. 3.1; Mismatches 1; Indels 0; Gaps 0;  
 Matches 9; Conservative 1; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
 |||||  
 Db 134 LLIDNKGVIKL 144

RESULT 7  
 B44349  
 protein kinase (EC 2.7.1.37) cdc2-B - African clawed frog  
 N:Alternate names: maturation-promoting factor p34cdc2 chain B  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 A:Accession: B44349  
 R:Pickham, K.M.; Meyer, A.N.; Li, J.; Donoghue, D.J.  
 Mol. Cell. Biol. 12, 3192-3203, 1992  
 A:Title: Requirement of mos(Xe) protein kinase for meiotic maturation of Xenopus oocytes  
 A:Reference number: A44349; MUID:92318937; PMID:1377775  
 A:Accession: B44349  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-302 <PIC>  
 A:Cross-references: UNIPROT:P24033; GB:M60682  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase  
 F:2-256/Domain: protein kinase homology <KIN>  
 F:10-18/Region: protein kinase ATP-binding motif  
 F:33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 81.1%; Score 43; DB 2; Length 302;  
 Best Local Similarity 81.8%; Pred. No. 3.1; Mismatches 1; Indels 0; Gaps 0;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
 |||||  
 Db 134 LLIDNKGVIKL 144

RESULT 8  
 A44349  
 protein kinase (EC 2.7.1.37) cdc2-A [similarity] - African clawed frog  
 N:Alternate names: maturation-promoting factor p34cdc2 chain A  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
 A:Accession: A44349  
 R:Pickham, K.M.; Meyer, A.N.; Li, J.; Donoghue, D.J.  
 Mol. Cell. Biol. 12, 3192-3203, 1992  
 A:Title: Requirement of mos(Xe) protein kinase for meiotic maturation of Xenopus oocytes  
 A:Reference number: A44349; MUID:92318937; PMID:1377775  
 A:Accession: A44349  
 A:Molecule type: mRNA  
 A:Residues: 1-302 <PIC>  
 A:Cross-references: UNIPROT:P35567; GB:M60680; NID:g2114022; PIDN:AAA63561.1; PID:g2114023  
 A:Experimental source: oocytes  
 A:Note: sequence extracted from NCBI backbone (NCBIN:107682, NCBIIP:107683)  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase  
 F:2-256/Domain: protein kinase homology <KIN>  
 F:10-18/Region: protein kinase ATP-binding motif  
 F:33,51,128,130/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 79.2%; Score 42; DB 1; Length 302;  
 Best Local Similarity 81.8%; Pred. No. 4.8; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
 |||||  
 Db 134 LLIDSKGVIKL 144

RESULT 9

T02306  
 probable protein kinase [imported] - Arabidopsis thaliana  
 N:Alternate names: hypothetical protein F13P17.2  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Aug-2004  
 A:Accession: T02306; D84753  
 R:Roundley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
 submitted to the EMBL Data Library, July 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC F13P17 genomic sequence.  
 A:Reference number: Z14657  
 A:Accession: T02306  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-502 <ROU>  
 A:Cross-references: UNIPROT:O22971; EMBL:AC004481; NID:g3337347; PIDN:AAC27394.1; PID:g3337348  
 A:Experimental source: cultivar Columbia  
 R:Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: D84753  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-502 <STO>  
 A:Cross-references: GB:AB002093; NID:g3337349; PIDN:AAC27394.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g34180; F13P17.2  
 A:Map position: 2  
 C:Superfamily: protein kinase homology  
 F:55-311/Domain: protein kinase homology <KIN>

Query Match 79.2%; Score 42; DB 2; Length 502;  
 Best Local Similarity 63.6%; Pred. No. 8; Mismatches 7; Conservative 3; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
 |||||  
 Db 185 LLDDKGNKV 195

RESULT 10  
 T07875  
 protein kinase homolog BNPK-3 - rape (fragment)  
 C:Species: Brassica napus (rape)  
 C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
 C:Accession: T07875  
 R:Zhou, X.R.  
 submitted to the EMBL Data Library, May 1995  
 A:Reference number: Z16189  
 A:Accession: T07875  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-69 <ZHO>  
 A:Cross-references: UNIPROT:Q39300; EMBL:U26747; NID:g849146; PIDN:AAB02871.1; PID:g849147  
 A:Experimental source: cv. H165; seedlings

Query Match 77.4%; Score 41; DB 2; Length 69;  
 Best Local Similarity 63.6%; Pred. No. 1.6; Mismatches 3; Indels 0; Gaps 0;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
 |||||  
 Db 10 LLDDKGNKV 20

RESULT 11  
 T47581  
 SCARECROW1 - Arabidopsis thaliana  
 N:Alternate names: protein P24B22.180  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C;Accession: T47581  
R;Bloecker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: Z23016  
A;Accession: T47581  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-653 <BLO>  
A;Cross-references: UNIPROT:Q9M384; EMBL:AL132957  
A;Experimental source: cultivar Columbia; BAC clone F24B22  
C;Genetics:  
A;Map position: 3  
A;Introns: 511/2  
A;Note: F24B22.180

Query Match 77.4%; Score 41; DB 2; Length 653;  
Best Local Similarity 70.0%; Pred. No. 16;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LIDDKGITKL 11  
|:|:|:|:|:  
Db 627 LVDDNGITKL 636

## RESULT 12

T51244  
SCARECROW protein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 28-Jul-2000 #sequence\_revision 28-Jul-2000 #text\_change 09-Jul-2004  
C;Accession: T51244  
R;di Laurenzio, L.; Wysocka-Diller, J.; Malamy, J.E.; Pysh, L.; Helariutta, Y.; Freshour  
submitted to the EMBL Data Library, July 1996  
A;Reference number: Z25338  
A;Accession: T51244  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-653 <DIL>  
A;Cross-references: UNIPROT:Q96304; EMBL:U62798; PIDN:AAB06318.1  
C;Genetics:  
A;Gene: SCARECROW1  
A;Introns: 511/2

Query Match 77.4%; Score 41; DB 2; Length 653;  
Best Local Similarity 70.0%; Pred. No. 16;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LIDDKGITKL 11  
|:|:|:|:|:  
Db 627 LVDDNGITKL 636

## RESULT 13

F82469  
hypothetical protein VCA0342 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C;Species: Vibrio cholerae  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: F82469  
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.  
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A;Reference number: A82035; MUID:20406833; PMID:10952301  
A;Accession: F82469  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-100 <HEI>  
A;Cross-references: UNIPROT:Q9KWK0; GB:AE004372; GB:AE003853; NID:g9657741; PIDN:AAF9625  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VCA0342  
A;Map position: 2

Query Match 75.5%; Score 40; DB 2; Length 100;  
Best Local Similarity 77.8%; Pred. No. 3.6;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 IDDKGTIKL 11  
|:|:|:|:|:  
Db 58 VDDKGTIQL 66

## RESULT 14

PQ0207  
protein kinase (EC 2.7.1.37) 1 - slime mold (Dictyostelium discoideum) (fragment)  
C;Species: Dictyostelium discoideum  
C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Aug-2004  
C;Accession: PQ0207  
R;Buerki, E.; Anjard, C.; Scholder, J.C.; Raymond, C.D.  
Gene 102, 57-65, 1991  
A;Title: Isolation of two genes encoding putative protein kinases regulated during Dicty  
A;Reference number: JQ1150; MUID:91323730; PMID:1864510  
A;Accession: PQ0207  
A;Molecule type: mRNA  
A;Residues: 1-355 <BUE>  
A;Cross-references: UNIPROT:P34100; GB:M38704  
C;Genetics:  
A;Gene: PK1  
C;Superfamily: protein kinase homology  
C;Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k  
F;2-258/Domain: protein kinase homology <KIN>  
F;10-18/Region: protein kinase ATP-binding motif  
F;301-355/Region: glutamine-rich

Query Match 75.5%; Score 40; DB 2; Length 355;  
Best Local Similarity 72.7%; Pred. No. 13;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
|:|:|:|:|:  
Db 133 ILIDKGHILK 143

## RESULT 15

S49752  
homeotic protein YOX1 - Yeast (Saccharomyces cerevisiae)  
N;Alternate names: protein YML027w  
C;Species: Saccharomyces cerevisiae  
C;Date: 13-Jan-1995 #sequence\_revision 10-Feb-1995 #text\_change 16-Aug-2004  
C;Accession: S49752; S33388  
R;Badcock, K.; Churcher, C.  
submitted to the EMBL Data Library, November 1994  
A;Reference number: S49741  
A;Accession: S49752  
A;Molecule type: DNA  
A;Residues: 1-385 <BAD>  
A;Cross-references: UNIPROT:P34161; EMBL:Z46659; NID:g575680; PIDN:CAA86628.1; PID:g5756  
R;Kaufmann, E.  
Chromosoma 102, 174-179, 1993  
A;Title: In vitro binding to the leucine tRNA gene identifies a novel yeast homeobox gen  
A;Reference number: S33388; MUID:9209080; PMID:8096171  
A;Accession: S33388  
A;Molecule type: DNA  
A;Residues: 1-290, 'OQLIIP' <KAU>  
A;Cross-references: EMBL:X62392; NID:g5501; PIDN:CAA44264.1; PID:g5502  
C;Genetics:  
A;Gene: SGD:YOX1  
A;Cross-references: SGD:S0004489; MIPS:YML027w  
A;Map position: 13L  
C;Superfamily: homeobox homology  
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;177-233/Domain: homeobox homology <HOX>

Query Match 75.5%; Score 40; DB 2; Length 385;  
Best Local Similarity 77.8%; Pred. No. 14;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;



Qy 3 IDDKGTIKL 11  
:|:|:|:|:|  
Db 38 LDDRGTIKL 46

Search completed: November 18, 2005, 01:09:41  
Job time : 11.9817 secs

**This Page Blank (usp10)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:04:26 ; Search time 48.8495 Seconds  
(without alignment)  
115.311 Million cell updates/sec

Title: US-10-006-177-9

Perfect score: 53

Sequence: 1 LLIDDKGKTIKL 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	100.0	79	2	Q9GKX0
2	53	100.0	191	2	Q8R4A4
3	53	100.0	295	1	Q9JW7
4	53	100.0	297	1	CDC2_BOVIN
5	53	100.0	297	1	CDC2_HUMAN
6	53	100.0	297	1	CDC2_MOUSE
7	53	100.0	297	1	CDC2_RAT
8	53	100.0	297	2	Q66GZ7
9	48	90.6	303	1	CDC2_CHICK
10	43	81.1	184	2	Q6QMT0
11	43	81.1	299	2	Q62572
12	43	81.1	300	2	Q17066
13	43	81.1	301	2	O18331
14	43	81.1	302	1	CC22_XENLA
15	43	81.1	302	1	CDC2_CARAU
16	43	81.1	302	2	Q7SZ44
17	43	81.1	302	2	Q7T3L7
18	43	81.1	303	1	CDC2_ORYCU
19	43	81.1	303	1	CDC2_ORYJA
20	43	81.1	303	1	CDC2_ORYLA
21	43	81.1	303	1	CDC2_ORYLU
22	43	81.1	488	2	Q7PYC4
23	43	81.1	488	2	Q42625
24	43	81.1	666	2	Q7RVV6
25	42	79.2	302	1	CC21_XENLA
26	42	79.2	302	1	CDC2_RANDY
27	42	79.2	302	2	Q6P7L3
28	42	79.2	502	2	O22971
29	41	77.4	69	2	Q39300
30	41	77.4	435	2	Q6M0B8
31	41	77.4	653	2	Q96304

32 41 77.4 653 2 Q9M384  
33 40 75.5 100 2 Q9KMK0  
34 40 75.5 230 2 Q8I584  
35 40 75.5 385 1 YOK1\_YEAST  
36 40 75.5 416 1 PKD1\_DICDI  
37 40 75.5 714 2 Q86AR2  
38 40 75.5 1025 2 Q9PQ11  
39 40 75.5 1641 2 Q8BLR9  
40 40 75.5 2197 2 Q74MS4  
41 39 73.6 223 2 Q6MZC2  
42 39 73.6 223 2 Q64DP9  
43 39 73.6 223 2 Q64DM7  
44 39 73.6 238 2 Q55I18  
45 39 73.6 268 2 Q7VI95

#### ALIGNMENTS

##### RESULT 1

Q9GKX0 PRELIMINARY; PRT; 79 AA.  
AC Q9GKX0;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Cdc2 protein (Fragment).  
GN Name=cdc2;  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Uchida T.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AB045783; BAB20908.1;  
DR GO; GO:000524; F:ATP binding; IEA.  
DR GO; GO:0004672; F:protein kinase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR011009; Kinase like.  
DR InterPro; IPR000719; Prot.kinase.  
DR Pfam; PF00069; Pkinase\_1.  
DR ProDom; PD000001; Prot.kinase; 1.  
DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
FT NON TER 1 1  
FT NON TER 79 79  
SQ SEQUENCE 79 AA; 9275 MW; E4C089237C1D4FFD CRC64;

Query Match 100.0%; Score 53; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 0.054;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGKTIKL 11

Db 66 LLIDDKGKTIKL 76

##### RESULT 2

Q8R4A4 PRELIMINARY; PRT; 191 AA.  
ID Q8R4A4;  
AC Q8R4A4;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Cell cycle p34 CDC2 kinase protein (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SvEvTacrBr;

```

RA Trickett J.I., Giese K.P.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: AF488732; AA09474.1; -.
DR HSP: P24941; 1H08.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER 1
SQ SEQUENCE 191 AA; 21920 MW; 4C906851F00701D4 CRC64;

Query Match 100.0%; Score 53; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 28 LLIDDKGTIKL 38

RESULT 3
Q99JW7 PRELIMINARY; PRT; 295 AA.
ID CDC2_BOVIN STANDARD; PRT; 297 AA.
AC P48734;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cdc2a protein (Fragment).
GN Name=Cdc2a;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RA STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalska U., Smallus D.B., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
SEQUENCE FROM N.A.
RA STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RC Strausberg R.;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: BC005614; AA05614.1; -.

```

```

DR HSP: P24941; 1O1O.
DR MGD; MGI.88351; Cdc2a.
DR GO: GO:0030544; F:Hsp70 protein binding; IPI.
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:0007095; P:mitotic G2 checkpoint; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER 1
SQ SEQUENCE 295 AA; 33832 MW; F6A2774A0154B9E1 CRC64;

Query Match 100.0%; Score 53; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 132 LLIDDKGTIKL 142

RESULT 4
CDC2_BOVIN STANDARD; PRT; 297 AA.
ID CDC2_BOVIN STANDARD; PRT; 297 AA.
AC P48734;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.37) (p34 protein
kinase) (Cyclin-dependent kinase 1) (CDK1).
GN Name=CDK2; Synonyms=CDK1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A.
RA MEDLINE=94215918; PubMed=8163203; DOI=10.1016/0378-1119(94)90586-X;
RX Yang L.I., Farin C.E.;
RT "Identification of cDNAs encoding bovine cyclin B and Cdk1/Cdc2.";
RL Gene 141:283-286 (1994).
CC -1- FUNCTION: Plays a key role in the control of the eukaryotic cell
cycle. It is required in higher cells for entry into S-phase and
mitosis. p34 is a component of the kinase complex that
phosphorylates the repetitive carboxyl-terminus of RNA polymerase
II.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- ENZYME REGULATION: Phosphorylation at Thr-14 or Tyr-15 inactivates
the enzyme, while phosphorylation at Thr-161 activates it (By
similarity).
CC -1- SUBUNIT: Forms a stable but non-covalent complex with a regulatory
subunit and with a cyclin.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC CDC2/CDKX subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

```

DR EMBL; L26547; AAA18894.1; -;  
 DR PIR; I45977; I45977.  
 DR HSGP; P24941; 1P2A.  
 DR InterPro; IPR011009; Kinase like.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR ATP-binding; Cell cycle; Cell division; Mitosis; Nuclear protein; Kinase; Transferase.  
 KW Phosphorylation; Serine/threonine-protein kinase; Mitosis; Nuclear protein; Kinase; Transferase.  
 FT DOMAIN 4 287  
 FT NP\_BIND 10 18 ATP (By similarity).  
 FT BINDING 33 33 ATP (By similarity).  
 FT ACT\_SITE 128 128 Proton acceptor (By similarity).  
 FT MOD\_RES 14 14 Phosphothreonine (By similarity).  
 FT MOD\_RES 15 15 Phosphotyrosine (By similarity).  
 FT MOD\_RES 161 161 Phosphothreonine (by CAK) (By similarity).  
 SQ SEQUENCE 297 AA; 33926 MW; 59F4D64BDESA79B CRC64;

Query Match 100.0%; Score 53; DB 1; Length 297;  
 Best Local Similarity 100.0%; Pred. No. 0.2;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
 Db 134 LLIDDKGTIKL 144

RESULT 5  
 CDC2\_HUMAN  
 ID\_CDC2\_HUMAN STANDARD; PRT; 297 AA.  
 AC P06493; O60764;  
 DT 01-JAN-1998 (Rel. 06, Created)  
 DT 01-JAN-1998 (Rel. 06, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Cell division control protein 2 homolog (EC 2.7.1.37) (p34 protein kinase) (Cyclin-dependent kinase 1) (CDK1).  
 DE Name=CDK2;  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 ON [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87201915; PubMed=3553962; DOI=10.1038/327031a0;  
 RA Lee M.G., Nurse P.;  
 RT "Complementation used to clone a human homologue of the fission yeast cell cycle control gene cdc2.";  
 RL Nature 327:31-35(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT CDC2DELTA.  
 RC TISSUE=Breast cancer;  
 RX MEDLINE=98175539; PubMed=9515786;  
 RA Ohta T., Okamoto K., Ishihashi F., Shibata K., Fukuda M., Yamaguchi S., Xiong Y.;  
 RT "T-loop deletion of CDC2 from breast cancer tissues eliminates binding to cyclin B1 and cyclin-dependent kinase inhibitor p21.";  
 RL Cancer Res. 58:1095-1098(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W., Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D., Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;  
 RT "NIHES-SNPs, environmental genome project, NIHES ES15478, Department of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T.I., Max S.I., Wang J., Hong L., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh F., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., C., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP PHOSPHORYLATION, AND ASSOCIATION WITH P13.  
 RX MEDLINE=88253421; PubMed=3289755; DOI=10.1016/0092-8674(88)90175-4;  
 RA Draetta G., Beach D.;  
 RT "Activation of cdc2 protein kinase during mitosis in human cells: cell cycle-dependent phosphorylation and subunit rearrangement.";  
 RL Cell 54:17-26(1988).  
 CC -!- FUNCTION: Plays a key role in the control of the eukaryotic cell cycle. It is required in higher cells for entry into S-phase and mitosis. p34 is a component of the kinase complex that phosphorylates the repetitive carboxyl-terminus of RNA polymerase II.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- ENZYME REGULATION: Phosphorylation at Thr-14 or Tyr-15 inactivates the enzyme, while phosphorylation at Thr-161 activates it.  
 CC -!- SUBUNIT: Forms a stable but non-covalent complex with a regulatory subunit and with a cyclin.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -!- POLYMORPHISM: Isoform CDC2delta1 is found in breast cancer tissues. It is unable to complex with cyclin B1 and also fails to bind to the CDK inhibitor p21.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. CDC2/CDK subfamily.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC EMBL; X05360; CAA28963.1; -;  
 CC EMBL; Y00272; CAA68376.1; -;  
 CC EMBL; D88357; BAA2600.1; -;  
 CC EMBL; AF512554; AAM34793.1; -;  
 CC EMBL; BC014563; AAH14563.1; -;  
 CC PIR; A29539; A29539.  
 CC PDB; 1LC9; Model; A=1-297.  
 CC SWISS-2DPAGE; P06493; HUMAN.  
 CC Genew; HGNC:1122; CDC2.  
 CC H-InvDB; HIX0008851; -;  
 CC Reactome; P06493; -;  
 CC MTM; 116940; -;  
 CC GO; GO:0005634; C:nucleus; TAS.  
 CC InterPro; IPR011009; Kinase like.  
 CC InterPro; IPR000719; Prot kinase.  
 CC InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 CC InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC Pfam; PF00069; Pkinase; 1.

```

DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S TKC_1
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR 3D-structure; ATP-binding; Cell cycle; Cell division; Mitosis;
KW Nucleic protein; Phosphorylation; Polymorphism;
KW Serine/threonine-protein kinase; Transferase.
FT DOMAIN 4 287
FT NP_BIND 10 18
FT BINDING 33 33 ATP (By similarity).
FT ACT_SITE 128 128 Proton acceptor (By similarity).
FT MOD_RES 14 14 Phosphothreonine (By similarity).
FT MOD_RES 15 15 Phosphotyrosine (By similarity).
FT MOD_RES 161 161 Phosphothreonine (By CAK) (By similarity).
FT VARIANT 107 163
FT VARIANT 107 163 Missing (in CDC2deltaT).
FT VARIANT 107 163 /FTID=VAR_011629.
SQ SEQUENCE 297 AA; 34095 MW; 942D79478FE490A CRC64;
Query Match 100.0%; Score 53; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. NO. 0.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144
RESULT 6
CDC2_MOUSE
ID _CDC2_MOUSE STANDARD; PRT; 297 AA.
AC P11440; P70337;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.37) (p34 protein
DE kinase) (Cyclin-dependent kinase 1) (CDK1).
GN Name=Cdc2; Synonyms=Cdc2a;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=FW3A; TISSUE=Breast carcinoma;
RX MEDLINE=91004239; PubMed=2209288; DOI=10.1016/0092-8674(90)90164-A;
RA Th'Ng J.P.H., Wright P.S., Hamaguchi J., Lee M.G., Norbury C.J.,
RA Nurse P., Bradbury E.M.;
RT "The FT210 cell line is a mouse G2 phase mutant with a temperature-
RT sensitive CDC2 gene product.";
RL Cell 63:313-324 (1990).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=92119319; PubMed=2132958;
RA Spurr N.K., Gough A.C., Lee M.G.;
RT "Cloning of the mouse homologue of the yeast cell cycle control gene
RT cdc2.";
RL DNA Seq. 1:49-54 (1990).
[3]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=89295557; PubMed=2662013; DOI=10.1038/339679a0;
RA Claeck L.J., Corden J.L.;
RT "Phosphorylation of RNA polymerase by the murine homologue of the
RT cell-cycle control protein cdc2.";
RL Nature 339:679-684 (1989).
[4]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Jun D., Park H., Lee Y., Kim Y.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
[5]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaado I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimonod S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.O., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
[6]
RN SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[7]
RN FUNCTION: Plays a key role in the control of the eukaryotic cell
RN cycle. It is required in higher cells for entry into S-phase and
RN mitosis. p34 is a component of the kinase complex that
RN phosphorylates the repetitive carboxyl-terminus of RNA polymerase
RN II.
[8]
RN CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
[9]
RN ENZYME REGULATION: Phosphorylation at Thr-14 or Tyr-15 inactivates
RN the enzyme, while phosphorylation at Thr-161 activates it (By
RN similarity).
[10]
RN SUBUNIT: Forms a stable but non-covalent complex with a regulatory
RN subunit and with a cyclin.
[11]
RN SUBCELLULAR LOCATION: Nuclear (By similarity).
[12]
RN SIMILARITY: Belongs to the Ser/Thr protein kinase family.
RN CDC2/CDKX subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; M38724; AAA37408.1; -.
CC EMBL; X16461; CAA34481.1; -.
CC EMBL; U58633; AAB09465.1; -.
CC EMBL; AK030231; BAC26856.1; -.
CC EMBL; BC024396; AAH24396.1; -.
CC PIR; A36074; A36074.
CC HSSP; P24941; 1H26.
CC MGP; MGI188351; Cdc2a.
CC GO; GO:0030544; F:Heb70 protein binding; IPI.
CC GO; GO:0016301; F:kinase activity; IDA.
CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
CC InterPro; IPR011009; Kinase like.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; Pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKC_1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC ATP-binding; Cell cycle; Cell division; Direct protein sequencing;
KW Mitosis; Nuclear protein; Phosphorylation;
KW Serine/threonine-protein kinase; Transferase.
FT DOMAIN 4 287
FT NP_BIND 10 18 ATP (By similarity).
FT BINDING 33 33 ATP (By similarity).
FT ACT_SITE 128 128 Proton acceptor (By similarity).
FT MOD_RES 14 14 Phosphothreonine (By similarity).
FT MOD_RES 15 15 Phosphotyrosine (By similarity).
FT MOD_RES 161 161 Phosphothreonine (by CAK) (By similarity).
FT CONFLICT 112 112 I -> M (in Ref. 4).
FT CONFLICT 113 113 L -> M (in Ref. 3).
FT CONFLICT 165 165 V -> L (in Ref. 2).
FT CONFLICT 245 245 K -> N (in Ref. 3 and 4).
FT CONFLICT 260 260 G -> C (in Ref. 3 and 4).
FT CONFLICT 263 263 L -> F (in Ref. 3 and 4).
FT CONFLICT 273 273 A -> T (in Ref. 2).
SQ SEQUENCE 297 AA; 34106 MW; 9C399FCC41BA7F31 CRC64;

Query Match 100.0%; Score 53; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144

RESULT 7
CDC2_RAT STANDARD; PRT; 297 AA.
AC P39951;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cell division control protein 2 homolog (EC 2.7.1.37) (p34 protein kinase) (Cyclin-dependent kinase 1) (CDK1).
GN Name=Cdc2; Synonyms=Cdc2a;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Kanaoka Y., Nojima H., Okayama H.;

```

```

RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SUBCELLULAR LOCATION.
RX MEDLINE=89249301; PubMed=2541912; DOI=10.1016/0092-8674(89)90914-8;
RA Rabinow K., Draetta G., Brizuela L., Vandre D., Beach D.;
RT "The cdc2 kinase is a nuclear protein that is essential for mitosis in mammalian cells";
RL Cell 57:393-401(1989).
CC CC -1- FUNCTION: Plays a key role in the control of the eukaryotic cell cycle. It is required in higher cells for entry into S-phase and mitosis. p34 is a component of the kinase complex that phosphorylates the repetitive carboxyl-terminus of RNA polymerase II.
CC CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC CC -1- ENZYME REGULATION: Phosphorylation at Thr-14 or Tyr-15 inactivates the enzyme, while phosphorylation at Thr-161 activates it (By similarity).
CC CC -1- SUBUNIT: Forms a stable but non-covalent complex with a regulatory subunit and with a cyclin.
CC CC -1- SUBCELLULAR LOCATION: Nuclear.
CC CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. CDC2/CDKX subfamily.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).


EMBL; X60767; CAA43177.1; -.
PIR; S24913; S24913.
HSSP; P24941; 1H26.
RGD; 2319; Cdc2a.
InterPro; IPR011009; Kinase like.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR008271; Ser_thr_pkin_AS.
InterPro; IPR002290; Ser_thr_pkinase.
Pfam; PF00069; Pkinase; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00220; S_TKC_1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
ATP-binding; Cell cycle; Cell division; Mitosis; Nuclear protein; Phosphorylation; Serine/threonine-protein kinase; Transferase.
FT DOMAIN 4 287
FT NP_BIND 10 18 ATP (By similarity).
FT BINDING 33 33 ATP (By similarity).
FT ACT_SITE 128 128 Proton acceptor (By similarity).
FT MOD_RES 14 14 Phosphothreonine (By similarity).
FT MOD_RES 15 15 Phosphotyrosine (By similarity).
FT MOD_RES 161 161 Phosphothreonine (by CAK) (By similarity).
SQ SEQUENCE 297 AA; 34134 MW; 73695017E127178A CRC64;

Query Match 100.0%; Score 53; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144

RESULT 8
Q66GZ7 PRELIMINARY; PRT; 297 AA.
AC Q66GZ7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)


```

```

DE MGC97947 protein.
GN Name=NGC97947;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX PubMed=1247932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.; to the EMBL/GenBank/DBJ databases.
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082118; AAH82118.1; -.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 297 AA; 34106 MW; 9C399FCC41BA7F31 CRC64;

Query Match 100.0%; Score 53; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.2; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

Qy 1 LLIDDKGTIKL 11
Db 134 LLIDDKGTIKL 144

RESULT 9
CDC2_CHICK
ID CDC2_CHICK STANDARD; PRT; 303 AA.
AC P13863;

```



```

FT  MOD_RES  277  277  Phosphoserine.
SQ  SEQUENCE  303 AA; 34688 MW; 976740ECC4741D69 CRC64;

Query Match      90.6%; Score 48; DB 1; Length 303;
Best Local Similarity 90.9%; Pred. No. 1.8;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1 LLIDDKGTIKL 11
Db  134 LLIDDKGVIKL 144

RESULT 10
O6QWTO PRELIMINARY; PRT; 184 AA.
ID  O6QWTO
AC  O6QWTO; 277, Created
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE  Cyclin-dependent kinase 1 (Fragment).
GN  Name=Cdc2;
OS  Anabas testudineus (Climbing perch).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC  Acantomorpha; Acanthopterygii; Percomorpha; Perciformes;
OC  Anabantoidae; Anabantidae; Anabas.
OX  NCBI_TaxID=64144;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  PubMed=15151934;
RX  Basu D., Navneet A.K., Dasgupta S., Bhattacharya S.;
RT  "CDC2-cyclin B-induced G2 to M transition in perch oocyte is dependent
RL  on Cdc25.";
CC  -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR  EMBL; AY533308; AAS59851.1; --
DR  HSSP; P24941; 1A01.
DR  GO; GO:0005524; F:ATP binding; IEA.
DR  GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR  GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR  GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR  InterPro; IPR011009; Kinase like.
DR  InterPro; IPR000719; Prot_kinase.
DR  InterPro; IPR008271; Ser_thr_pkinase.
DR  InterPro; IPR01245; Tyr_pkinase.
DR  ProDom; PD000001; Prot_kinase; 1.
DR  SMART; SM00220; S_TKC; 1.
DR  SMART; SM00219; TyrKC; 1.
DR  PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR  PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
KW  ATP-binding; Cyclin; Kinase; Serine/threonine-protein kinase;
KW  Transferase.
FT  NON_TER  1
FT  NON_TER  184
SQ  SEQUENCE  184 AA; 21186 MW; F788083CC868BBC4 CRC64;

Query Match      81.1%; Score 43; DB 2; Length 184;
Best Local Similarity 81.8%; Pred. No. 9;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy  1 LLIDDKGTIKL 11
Db  106 LLIDDKGVIKL 116

RESULT 11
O62572 PRELIMINARY; PRT; 299 AA.
ID  O62572;
AC  O62572;
DT  01-AUG-1998 (TrEMBLrel. 07, Created)
DT  01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

```

DE  Cyclin dependent kinase 1.
OS  Sphaerechinus granularis (Purple sea urchin).
OC  Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC  Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC  Sphaerechinus.
OX  NCBI_TaxID=39374;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  TISSUE=Whole embryo;
RA  Moreau J.L., Marques F., Barakat A., Schatt P., Lozano J.C.,
RA  Peaucellier G., Picard A., Genevriere A.M.;
RL  Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC  -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR  EMBL; AJ225013; CAA12343.1; --
DR  HSSP; P24941; 1OIQ.
DR  GO; GO:0005524; F:ATP binding; IEA.
DR  GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR  GO; GO:0016740; F:transferase activity; IEA.
DR  GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR  InterPro; IPR011009; Kinase like.
DR  InterPro; IPR000719; Prot_kinase.
DR  InterPro; IPR002290; Ser_thr_pkinase.
DR  InterPro; IPR008271; Ser_thr_pkin_AS.
DR  Pfam; PF00069; Pkinase; 1.
DR  SMART; SM00220; S_TKC; 1.
DR  PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR  PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR  PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
KW  ATP-binding; Cyclin; Kinase; Serine/threonine-protein kinase;
KW  Transferase.
SQ  SEQUENCE  299 AA; 34635 MW; 6E0FE48F6635FAA1 CRC64;

Query Match      81.1%; Score 43; DB 2; Length 299;
Best Local Similarity 81.8%; Pred. No. 15;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy  1 LLIDDKGTIKL 11
Db  133 LLIDDKGTIKL 143

RESULT 12
Q17066 PRELIMINARY; PRT; 300 AA.
ID  Q17066
AC  Q17066;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Cdc2.
OS  Asterina pectinifera (Starfish).
OC  Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC  Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
OX  NCBI_TaxID=7594;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Okumura E., Sekiai T., Hisanaga S., Tachibana K., Kishimoto T.;
RA  "Initial Triggering of M-phase in Starfish Oocytes: A Possible Novel
RT  Component of Maturation-promoting Factor.";
RL  J. Cell Biol. 0:0-0(1996).
CC  -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR  EMBL; D79982; BAA11477.1; --
DR  HSSP; P24941; 1OIQ.
DR  GO; GO:0005524; F:ATP binding; IEA.
DR  GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR  GO; GO:0016740; F:transferase activity; IEA.
DR  GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR  InterPro; IPR011009; Kinase like.
DR  InterPro; IPR000719; Prot_kinase.
DR  InterPro; IPR002290; Ser_thr_pkinase.
DR  InterPro; IPR008271; Ser_thr_pkin_AS.
DR  Pfam; PF00069; Pkinase; 1.
DR  SMART; SM00220; S_TKC; 1.
DR  PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

```

```

DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 300 AA; 34209 MW; 3FCD5F78D5000FDF CRC64;

Query Match 81.1%; Score 43; DB 2; Length 300;
Best Local Similarity 81.8%; Pred. No. 15;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 133 LLIDDKGTIKL 143

RESULT 13
O18331
ID O18331 PRELIMINARY; PRT; 301 AA.
AC O18331; 1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P34cdc2.
OS Hemientrotus pulcherrimus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
OC Hemientrotus.
OX NCBI_TaxID=7650;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=98036069; PubMed=9370302; DOI=10.1016/S0378-1119(97)00338-7;
RA Komatsu S., Murata-Hori M., Totsukawa G., Murai N., Fujimoto H.,
RA Mabuchi I., Hosoya H.;
RT Identification of p34cdc2 kinase from sea urchin Hemientrotus
RT pulcherrimus and its involvement in the phosphorylation of myosin II
RT regulatory light chain in the metaphase extract."
RL Gene 198.359-365(1997).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; D82878; BAA23218.1; -.
DR HSP; P24941; 10IQ.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 301 AA; 34751 MW; AD2253E7230A72C6 CRC64;

Query Match 81.1%; Score 43; DB 2; Length 301;
Best Local Similarity 81.8%; Pred. No. 15;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 133 LLIDDKGTIKL 143

RESULT 14
CC22 XENLA
ID CC22 XENLA STANDARD; PRT; 302 AA.
AC P24033; 1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cell division control protein 2 homolog 2 (EC 2.7.1.37) (p34 protein
DE kinase 2).

```

```

GN Name=CDC2X1.2;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
SEQUENCE FROM N.A.
RP SEQUENCE Oocyte;
RC MEDLINE=92318937; PubMed=1377775;
RX Pickham K.M., Meyer A.N., Li J., Donoghue D.J.;
RA "Requirement of mosXe protein kinase for meiotic maturation of Xenopus
RT oocytes induced by a cdc2 mutant lacking regulatory phosphorylation
RT sites."
RL Mol. Cell. Biol. 12:3192-3203(1992).
RN [2]
PROPHORYLATION SITE THR-161.
RX MEDLINE=93345456; PubMed=8344251;
RA Pesquet D., Labbe J.-C., Derancourt J., Capony J.-P., Galas S.,
RA Girard F., Lorca T., Shuttleworth J., Dorée M., Cavadore J.-C.;
RT "The MO15 gene encodes the catalytic subunit of a protein kinase that
RT activates cdc2 and other cyclin-dependent kinases (CDKs) through
RT phosphorylation of Thr161 and its homologues."
RL EMBO J. 12:3111-3121(1993).
CC -1- FUNCTION: Plays a key role in the control of the eukaryotic cell
CC cycle. It is required in higher cells for entry into S-phase and
CC mitosis. Component of the kinase complex that phosphorylates the
CC repetitive carboxyl-terminus of RNA polymerase II (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- ENZYME REGULATION: Phosphorylation at Thr-14 or Tyr-15 inactivates
CC the enzyme, while phosphorylation at Thr-161 activates it.
CC -1- SUBUNIT: Forms a stable but non-covalent complex with a regulatory
CC subunit and with a cyclin.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC CDC2/CDKX subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M60681; AAA63562.1; -.
DR PIR; B43349; B43349.
DR HSP; P24941; 1P2A.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Cell cycle; Cell division; Mitosis; Nuclear protein;
KW Phosphorylation; Serine/threonine-protein kinase; Transferase.
FT DOMAIN 4 287
FT NP_BIND 10 18 ATP (By similarity).
FT BINDING 33 33 Proton acceptor (By similarity).
FT ACT_SITE 128 128 Phosphothreonine (By similarity).
FT MOD_RES 14 14 Phosphotyrosine (By similarity).
FT MOD_RES 15 15 Phosphothreonine (By similarity).
FT MOD_RES 161 161 Phosphoserine (By similarity).
FT MOD_RES 277 277 Phosphothreonine (By similarity).
SQ SEQUENCE 302 AA; 34429 MW; 42ED8507254282B3 CRC64;

Query Match 81.1%; Score 43; DB 1; Length 302;
Best Local Similarity 81.8%; Pred. No. 15;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 LLIDDKGTIKL 11  
 Db 134 LLIDNKGVIKL 144  
 RESULT 15  
 CDC2\_CARAU STANDARD; PRT; 302 AA.  
 AC PS1958;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Cell division control protein 2 homolog (EC 2.7.1.37) (p34 protein kinase) (Cyclin-dependent kinase 1) (CDK1).  
 GN Name=CDK2;  
 OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Carassius.  
 OK NCBI\_TaxID=7957;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Oocyte;  
 RA Kajura H., Yamashita M., Katsu Y., Nagahama Y.;  
 RT "Isolation and characterization of goldfish cdc2, a catalytic component of maturation-promoting factor.";  
 RL Dev. Growth Differ. 35:647-654(1993).  
 CC -!- FUNCTION: Plays a key role in the control of the eukaryotic cell cycle. It is required in higher cells for entry into S-phase and mitosis. Component of the kinase complex that phosphorylates the repetitive carboxyl-terminus of RNA polymerase II. Catalytic component of MPF.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- ENZYME REGULATION: Phosphorylation at Thr-14 or Tyr-15 inactivates the enzyme, while phosphorylation at Thr-161 activates it (By similarity).  
 CC -!- SUBUNIT: Forms a stable but non-covalent complex with cyclin B in mature oocytes.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 CC CDC2/CDKX subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D17758; BAA04605.1; -;  
 CC PIR; I50474; I50474.  
 CC HSSP; P24941; 1P2A.  
 CC InterPro; IPR011009; Kinase like.  
 CC InterPro; IPR000719; Prot\_kinase.  
 CC InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 CC InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC Pfam; PF00069; Pkinase; I.  
 CC ProDom; PD000001; Prot\_kinase; 1.  
 CC SMART; SM00220; S\_TKC; 1.  
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 CC PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 CC ATP-binding; Cell cycle; Cell division; Mitosis; Nuclear protein;  
 KW Phosphorylation; Serine/threonine-protein kinase; Transferase.  
 FT DOMAIN 4 287 Protein kinase.  
 FT NP\_BIND 10 18 ATP (By similarity).  
 FT BINDING 33 33 ATP (By similarity).  
 FT ACT\_SITE 128 128 Proton acceptor (By similarity).  
 FT MOD\_RES 14 14 Phosphothreonine (By similarity).  
 FT MOD\_RES 15 15 Phosphotyrosine (By similarity).  
 FT MOD\_RES 161 161 Phosphothreonine (by CAK) (By

QY 1 LLIDDKGTIKL 11  
 Db 134 LLIDNKGVIKL 144  
 RESULT 15  
 CDC2\_CARAU STANDARD; PRT; 302 AA.  
 AC PS1958;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Cell division control protein 2 homolog (EC 2.7.1.37) (p34 protein kinase) (Cyclin-dependent kinase 1) (CDK1).  
 GN Name=CDK2;  
 OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Carassius.  
 OK NCBI\_TaxID=7957;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Oocyte;  
 RA Kajura H., Yamashita M., Katsu Y., Nagahama Y.;  
 RT "Isolation and characterization of goldfish cdc2, a catalytic component of maturation-promoting factor.";  
 RL Dev. Growth Differ. 35:647-654(1993).  
 CC -!- FUNCTION: Plays a key role in the control of the eukaryotic cell cycle. It is required in higher cells for entry into S-phase and mitosis. Component of the kinase complex that phosphorylates the repetitive carboxyl-terminus of RNA polymerase II. Catalytic component of MPF.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- ENZYME REGULATION: Phosphorylation at Thr-14 or Tyr-15 inactivates the enzyme, while phosphorylation at Thr-161 activates it (By similarity).  
 CC -!- SUBUNIT: Forms a stable but non-covalent complex with cyclin B in mature oocytes.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 CC CDC2/CDKX subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D17758; BAA04605.1; -;  
 CC PIR; I50474; I50474.  
 CC HSSP; P24941; 1P2A.  
 CC InterPro; IPR011009; Kinase like.  
 CC InterPro; IPR000719; Prot\_kinase.  
 CC InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 CC InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC Pfam; PF00069; Pkinase; I.  
 CC ProDom; PD000001; Prot\_kinase; 1.  
 CC SMART; SM00220; S\_TKC; 1.  
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 CC PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 CC ATP-binding; Cell cycle; Cell division; Mitosis; Nuclear protein;  
 KW Phosphorylation; Serine/threonine-protein kinase; Transferase.  
 FT DOMAIN 4 287 Protein kinase.  
 FT NP\_BIND 10 18 ATP (By similarity).  
 FT BINDING 33 33 ATP (By similarity).  
 FT ACT\_SITE 128 128 Proton acceptor (By similarity).  
 FT MOD\_RES 14 14 Phosphothreonine (By similarity).  
 FT MOD\_RES 15 15 Phosphotyrosine (By similarity).  
 FT MOD\_RES 161 161 Phosphothreonine (by CAK) (By

Search completed: November 18, 2005, 01:08:21  
 Job time : 51.8495 secs

QY 1 LLIDDKGTIKL 11  
 Db 134 LLIDNKGVIKL 144  
 RESULT 15  
 CDC2\_CARAU STANDARD; PRT; 302 AA.  
 AC PS1958;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Cell division control protein 2 homolog (EC 2.7.1.37) (p34 protein kinase) (Cyclin-dependent kinase 1) (CDK1).  
 GN Name=CDK2;  
 OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Carassius.  
 OK NCBI\_TaxID=7957;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Oocyte;  
 RA Kajura H., Yamashita M., Katsu Y., Nagahama Y.;  
 RT "Isolation and characterization of goldfish cdc2, a catalytic component of maturation-promoting factor.";  
 RL Dev. Growth Differ. 35:647-654(1993).  
 CC -!- FUNCTION: Plays a key role in the control of the eukaryotic cell cycle. It is required in higher cells for entry into S-phase and mitosis. Component of the kinase complex that phosphorylates the repetitive carboxyl-terminus of RNA polymerase II. Catalytic component of MPF.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- ENZYME REGULATION: Phosphorylation at Thr-14 or Tyr-15 inactivates the enzyme, while phosphorylation at Thr-161 activates it (By similarity).  
 CC -!- SUBUNIT: Forms a stable but non-covalent complex with cyclin B in mature oocytes.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 CC CDC2/CDKX subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D17758; BAA04605.1; -;  
 CC PIR; I50474; I50474.  
 CC HSSP; P24941; 1P2A.  
 CC InterPro; IPR011009; Kinase like.  
 CC InterPro; IPR000719; Prot\_kinase.  
 CC InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 CC InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC Pfam; PF00069; Pkinase; I.  
 CC ProDom; PD000001; Prot\_kinase; 1.  
 CC SMART; SM00220; S\_TKC; 1.  
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 CC PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 CC ATP-binding; Cell cycle; Cell division; Mitosis; Nuclear protein;  
 KW Phosphorylation; Serine/threonine-protein kinase; Transferase.  
 FT DOMAIN 4 287 Protein kinase.  
 FT NP\_BIND 10 18 ATP (By similarity).  
 FT BINDING 33 33 ATP (By similarity).  
 FT ACT\_SITE 128 128 Proton acceptor (By similarity).  
 FT MOD\_RES 14 14 Phosphothreonine (By similarity).  
 FT MOD\_RES 15 15 Phosphotyrosine (By similarity).  
 FT MOD\_RES 161 161 Phosphothreonine (by CAK) (By

Query Match 81.1%; Score 43; DB 1; Length 302;  
 Best Local Similarity 81.8%; Pred. No. 15;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
 Db 134 LLIDNKGVIKL 144

**This Page Blank (usp10)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 00:48:52 ; Search time 16.6774 Seconds  
(without alignments)  
49.237 Million cell updates/sec

Title: US-10-006-177-9

Perfect score: 53

Sequence: 1 LLIDDKGTIKL 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	100.0	82	1	US-08-446-038B-14
2	53	100.0	82	1	US-08-446-010B-14
3	53	100.0	82	2	US-08-805-445-14
4	53	100.0	82	2	US-08-064-067D-14
5	53	100.0	82	2	US-09-066-208-14
6	53	100.0	84	1	US-08-370-225-15
7	53	100.0	84	1	US-08-461-859-15
8	53	100.0	84	5	PCT-US93-10069-15
9	53	100.0	270	2	US-07-857-224B-31
10	53	100.0	297	1	US-08-176-620A-16
11	53	100.0	297	2	US-08-874-347-23
12	53	100.0	297	2	US-08-874-347-24
13	53	100.0	297	2	US-08-461-985-16
14	53	100.0	297	3	US-08-093-522-23
15	53	100.0	297	3	US-08-093-522-24
16	53	100.0	297	3	US-08-932-787B-21
17	53	100.0	297	3	US-08-932-012C-21
18	53	100.0	297	3	US-08-888-818C-21
19	53	100.0	297	4	US-09-411-628-12
20	53	100.0	297	4	US-10-174-794-12
21	53	100.0	299	4	US-09-949-016-10673
22	41	77.4	299	4	US-09-265-585C-103
23	41	77.4	368	4	US-09-265-585C-124
24	41	77.4	653	3	US-09-186-276B-2
25	41	77.4	653	4	US-08-842-445-2
26	41	77.4	653	4	US-08-186-188B-2
27	41	77.4	653	4	US-09-265-585C-2

#### ALIGNMENTS

##### RESULT 1

US-08-446-038B-14  
; Sequence 14, Application US/08446038B  
; Patent No. 5658791  
; GENERAL INFORMATION:  
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;  
; APPLICANT: Harpur, Ailsa  
; TITLE OF INVENTION: No. 5658791e1 Protein Tyrosine Kinase  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,038B  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/064,067  
; FILING DATE: 30-Jun-1993  
; APPLICATION NUMBER: PCT/US91/08889  
; FILING DATE: 26-No. 5658791-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: Australian PK3594/90  
; FILING DATE: 28-No. 5658791-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: Australian 88229/91  
; FILING DATE: 27-No. 5658791-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5658791man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5244  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-688-9200  
; TELEFAX: 212-838-3884  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 82 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-446-038B-14

Sequence 100, Appl  
Sequence 2121, Appl  
Sequence 10, Appl  
Sequence 8, Appl  
Sequence 9, Appl  
Sequence 21, Appl  
Sequence 6, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 20512, A  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 22616, A  
Sequence 6, Appl  
Sequence 18495, A  
Sequence 7, Appl  
Sequence 19, Appl

```
Query Match      100.0%; Score 53; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      18 LLIDDKGTIKL 28

RESULT 2
US-08-446-010B-14
; Sequence 14, Application US/08446010B
; Patent No. 5716818
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5716818el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,010B
; FILING DATE: 19-May-1995
; CLASSIFICATION: 433
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,038
; FILING DATE: 19-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-Jun-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5716818-1991
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5716818-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Baer, Madeline F.
; REGISTRATION NUMBER: 36,437
; REFERENCE/DOCKET NUMBER: LUD 5244.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-3884
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 82 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-446-010B-14

Query Match      100.0%; Score 53; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      18 LLIDDKGTIKL 28

RESULT 3
US-08-446-010B-14
; Sequence 14, Application US/08446010B
; Patent No. 5716818
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5716818el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,010B
; FILING DATE: 19-May-1995
; CLASSIFICATION: 433
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,038
; FILING DATE: 19-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-Jun-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5716818-1991
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5716818-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Baer, Madeline F.
; REGISTRATION NUMBER: 36,437
; REFERENCE/DOCKET NUMBER: LUD 5244.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-3884
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 82 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-446-010B-14

Query Match      100.0%; Score 53; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      18 LLIDDKGTIKL 28

RESULT 4
US-08-064-067D-14
; Sequence 14, Application US/08064067D
; Patent No. 5821069
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5821069el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,445
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,038
; FILING DATE: 19-MAY-1995
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-Jun-1993
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5821069-1991
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5821069-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian 88229/91
; FILING DATE: 27-No. 5821069-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5821069man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 82 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-805-445-14

Query Match      100.0%; Score 53; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      18 LLIDDKGTIKL 28

RESULT 4
US-08-064-067D-14
; Sequence 14, Application US/08064067D
; Patent No. 5821069
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
; APPLICANT: Harpur, Ailsa
; TITLE OF INVENTION: No. 5821069el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,445
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,038
; FILING DATE: 19-MAY-1995
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-Jun-1993
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5821069-1991
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5821069-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian 88229/91
; FILING DATE: 27-No. 5821069-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5821069man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 82 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-805-445-14
```

STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/064,067D  
FILING DATE: 30-Jun-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-No. 5852184-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-No. 5852184-1990  
APPLICATION NUMBER: 27-No. 5852184-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5852184man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELEPHONE: 212-838-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 82 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-064-067D-14

Query Match 100.0%; Score 53; DB 2; Length 82;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
| | | | | | | | | |  
Db 18 LLIDDKGTIKL 28

RESULT 5  
US-09-066-208-14  
Sequence 14, Application US/09066208  
Patent No. 5910426  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemicieki, Andrew;  
APPLICANT: Harpur, Alisa  
TITLE OF INVENTION: No. 5910426el Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/066,208  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/805,445  
FILING DATE: 25-FEB-1997  
APPLICATION NUMBER: US 08/446,038  
FILING DATE: 19-MAY-1995

APPLICATION NUMBER: 08/064,067  
FILING DATE: 30-Jun-1993  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-No. 5910426-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-No. 5910426-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-No. 5910426-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5910426man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELEPHONE: 212-838-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 82 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-066-208-14

Query Match 100.0%; Score 53; DB 2; Length 82;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
| | | | | | | | | |  
Db 18 LLIDDKGTIKL 28

RESULT 6  
US-08-370-225-15  
Sequence 15, Application US/08370225  
Patent No. 5580736  
GENERAL INFORMATION:  
APPLICANT: Brent, Roger  
APPLICANT: Gyuris, Jeno  
APPLICANT: Golemis, Erica  
TITLE OF INVENTION: Interaction Trap System for Isolating  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: Wordperfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/370,225  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/969,038  
FILING DATE: 10/30/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/143001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:

```
; LENGTH: 84
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-370-225-15
Query Match 100.0%; Score 53; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 61 LLIDDKGTIKL 71

RESULT 7
US-08-461-859-15
; Sequence 15, Application US/08461859
; Patent No. 5786169
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: Gyuris, Jenö
; APPLICANT: Golemis, Erica
; TITLE OF INVENTION: Interaction Trap System for Isolating
; TITLE OF INVENTION: Novel Proteins
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461.859
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,225
; FILING DATE: January 9, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/969,038
; FILING DATE: October 30, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/143002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-8906
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-461-859-15
Query Match 100.0%; Score 53; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 61 LLIDDKGTIKL 71

RESULT 8
PCT-US93-10069-15
; Sequence 15, Application PC/TUS9310069
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: Gyuris, Jenö
; APPLICANT: Golemis, Erica
; TITLE OF INVENTION: Interaction Trap System for Isolating
; TITLE OF INVENTION: Novel Proteins
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10069
; FILING DATE: 20-OCT-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/969,038
; FILING DATE: 10/30/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/143001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-10069-15
Query Match 100.0%; Score 53; DB 5; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 61 LLIDDKGTIKL 71

RESULT 9
US-07-857-224B-31
; Sequence 31, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
```

```
; LENGTH: 84
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-10-006-177-9.ra1
Query Match 100.0%; Score 53; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 61 LLIDDKGTIKL 71

RESULT 8
PCT-US93-10069-15
; Sequence 15, Application PC/TUS9310069
; GENERAL INFORMATION:
; APPLICANT: Brent, Roger
; APPLICANT: Gyuris, Jenö
; APPLICANT: Golemis, Erica
; TITLE OF INVENTION: Interaction Trap System for Isolating
; TITLE OF INVENTION: Novel Proteins
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10069
; FILING DATE: 20-OCT-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/969,038
; FILING DATE: 10/30/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/143001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-10069-15
Query Match 100.0%; Score 53; DB 5; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11
Db 61 LLIDDKGTIKL 71

RESULT 9
US-07-857-224B-31
; Sequence 31, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
```



SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA: US/07/857,224B  
FILING DATE: 03/25/92  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA: none  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (International) 41 1 632 2830  
TELEFAX: (International) 41 1 262 2437  
TELEX: none  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 270  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: protein  
ORIGINAL SOURCE: human  
ORGANISM: human  
FEATURE: Protein kinase; Table 8 Column 35  
PUBLICATION INFORMATION:  
AUTHORS:  
AUTHORS: Hanks, S. K.  
AUTHORS: Quinn, A. M.  
AUTHORS: Hunter, T.  
TITLE: The protein kinase family  
JOURNAL: Science  
VOLUME: 241  
PAGES: 42-52  
DATE: 1988  
US-07-857-224B-31

Query Match 100.0%; Score 53; DB 2; Length 270;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
| | | | | | | | | | | | | |  
Db 133 LLIDDKGTIKL 143

RESULT 10  
US-08-176-620A-16  
Sequence 16, Application US/08176620A  
Patent No. 5595904  
GENERAL INFORMATION:  
APPLICANT: Boulton, Teri G.  
APPLICANT: Cobb, Melanie H.  
APPLICANT: Yancopoulos, George D.  
APPLICANT: Nye, Steven  
APPLICANT: Panayotatos, Nikos  
TITLE OF INVENTION: A Family of Map2 Protein Kinases  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/176,620A  
FILING DATE: 03-JAN-1994  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Mirock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 6526-123

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 297 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-176-620A-16

Query Match 100.0%; Score 53; DB 1; Length 297;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
| | | | | | | | | | | | | |  
Db 134 LLIDDKGTIKL 144

RESULT 11  
US-08-874-347-23  
Sequence 23, Application US/08874347  
Patent No. 5863741  
GENERAL INFORMATION:  
APPLICANT: Limper, Andrew H.  
APPLICANT: Leof, Edward B.  
APPLICANT: Thomas, Charles F.  
APPLICANT: Gustafson, Michael P.  
TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
TITLE OF INVENTION: CARINII  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C., P.A.  
STREET: 60 South Sixth Street, Suite 3300  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/874,347  
FILING DATE: 13-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 07039/055001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 297 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-874-347-23

Query Match 100.0%; Score 53; DB 2; Length 297;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
| | | | | | | | | | | | | |

```
Db      134 LLIDDKGTIKL 144
;
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,985
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,620
; FILING DATE: 03-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-461-985-16
;
; Query Match 100.0%; Score 53; DB 2; Length 297;
; Best Local Similarity 100.0%; Pred. No. 0.021;
; Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy      1 LLIDDKGTIKL 11
;
; |||||
;
Db      134 LLIDDKGTIKL 144
;
; RESULT 14
; US-09-093-522-23
; Sequence 23, Application US/09093522
; Patent No. 6015700
; GENERAL INFORMATION:
; APPLICANT: Limper, Andrew H.
; APPLICANT: Leof, Edward B.
; APPLICANT: Thomas, Charles F.
; APPLICANT: Gustafson, Michael P.
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C., P.A.
; STREET: 60 South Sixth Street, Suite 3300
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,347
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ellinger, Mark S.
; REGISTRATION NUMBER: 34,812
; REFERENCE/DOCKET NUMBER: 07039/055001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-335-5070
; TELEFAX: 612-288-9696
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-874-347-24
;
; Query Match 100.0%; Score 53; DB 2; Length 297;
; Best Local Similarity 100.0%; Pred. No. 0.021;
; Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy      1 LLIDDKGTIKL 11
;
; |||||
;
Db      134 LLIDDKGTIKL 144
;
; RESULT 13
; US-08-461-985-16
; Sequence 16, Application US/08461985
; Patent No. 5872006
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
```

NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 07039/055002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 297 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-093-522-23

Query Match 100.0%; Score 53; DB 3; Length 297;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
|||  
Db 134 LLIDDKGTIKL 144

RESULT 15  
US-09-093-522-24  
; Sequence 24, Application US/09093522  
; Patent No. 6015700  
; GENERAL INFORMATION:  
; APPLICANT: Limber, Andrew H.  
; APPLICANT: Leof, Edward B.  
; APPLICANT: Thomas, Charles F.  
; APPLICANT: Gustafson, Michael P.  
; TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
; TITLE OF INVENTION: CARINII  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C., P.A.  
; STREET: 60 South Sixth Street, Suite 3300  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/093,522  
; FILING DATE: 08-JUN-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/874,347  
; FILING DATE: 13-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ellinger, Mark S.  
; REGISTRATION NUMBER: 34,812  
; REFERENCE/DOCKET NUMBER: 07039/055002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-335-5070  
; TELEFAX: 612-288-9696  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 297 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-093-522-24

Query Match 100.0%; Score 53; DB 3; Length 297;  
Best Local Similarity 100.0%; Pred. No. 0.021;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLIDDKGTIKL 11  
|||  
Db 134 LLIDDKGTIKL 144

Search completed: November 18, 2005, 09:36:02  
Job time : 17.6774 secs

**This Page Blank (uspto)**

Result No.	Score	Query Match	Length	DB	ID	Description
1	53	100.0	11	14	US-10-006-177-9	Sequence 9, Appl
2	53	100.0	297	14	US-10-060-065-11	Sequence 11, Appl
3	53	100.0	297	14	US-10-059-585-32	Sequence 32, Appl
4	53	100.0	297	14	US-10-177-293-41	Sequence 41, Appl
5	53	100.0	297	14	US-10-174-794-12	Sequence 12, Appl
6	53	100.0	297	14	US-10-204-041-8	Sequence 8, Appl
7	53	100.0	297	15	US-10-394-322A-6	Sequence 6, Appl
8	53	100.0	297	16	US-10-788-792-155	Sequence 155, App
9	53	100.0	297	16	US-10-733-878-419	Sequence 419, App
10	53	100.0	297	16	US-10-723-860-2448	Sequence 2448, Ap
11	53	100.0	297	16	US-10-751-736-67	Sequence 67, Appl

```
RESULT 2
US-10-060-065-11
; Sequence 11, Application US/10060065
; Publication No. US20030017480A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/10/060,065
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-065-11

Query Match      100.0%; Score 53; DB 14; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144

RESULT 3
US-10-059-585-32
; Sequence 32, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: Isogai, Takao
; APPLICANT: Nishikawa, Tetsuo
; APPLICANT: Hayashi, Koji
; APPLICANT: Otsuka, Kaoru
; APPLICANT: Yamamoto, Jun-ichi
; APPLICANT: Ishii, Shizuko
; APPLICANT: Sugiyama, Tomoyasu
; APPLICANT: Wakamatsu, Ai
; APPLICANT: Nagai, Keiichi
; APPLICANT: Otsuki, Tetsuji
; APPLICANT: Funahashi, Shin-ichi
```

```
; APPLICANT: Senoo, Chiaki
; APPLICANT: Nezu, Jun-ichi
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
; TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-098001
; CURRENT APPLICATION NUMBER: US/10/059,585
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05060
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-059-585-32

Query Match      100.0%; Score 53; DB 14; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144

RESULT 4
US-10-177-293-41
; Sequence 41, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: WPI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
```

```
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-41

Query Match      100.0%; Score 53; DB 14; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144

RESULT 5
US-10-174-794-12
; Sequence 12, Application US/10174794
; Publication No. US20030166220A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; FILE REFERENCE: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/10/174,794
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/411,628
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/102,906
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-794-12

Query Match      100.0%; Score 53; DB 14; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144

RESULT 6
US-10-204-041-8
; Sequence 8, Application US/10204041
; Publication No. US20030176443A1
; GENERAL INFORMATION:
; APPLICANT: STEIN-GERLACH, MATTHIAS
; APPLICANT: SALASSIDIS, KONSTADINOS
; APPLICANT: BACHER, GERALD
; APPLICANT: MULLER, STEFAN
; TITLE OF INVENTION: Pyridylpyrimidine Derivatives as Effective Compounds Against Pric
; FILE REFERENCE: Infections and Pric Diseases
; FILE REFERENCE: AXM-007.1P US
; CURRENT APPLICATION NUMBER: US/10/204,041
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: EP 0111858.5
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: PCT/EP02/05420
; PRIOR FILING DATE: 2002-05-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
```

```
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-204-041-8

Query Match      100.0%; Score 53; DB 14; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144

RESULT 7
US-10-394-322A-6
; Sequence 6, Application US/10394322A
; Publication No. US2003023291A1
; GENERAL INFORMATION:
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.
; APPLICANT: Prescott, John C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-322A-6

Query Match      100.0%; Score 53; DB 15; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144

RESULT 8
US-10-788-792-155
; Sequence 155, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: US/10/788,792
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 155
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-788-792-155

Query Match      100.0%; Score 53; DB 16; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144
```

```
RESULT 9
US-10-733-878-419
; Sequence 419, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: CHEMOKINES AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: THAP PROTEINS AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10733.878
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 419
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-733-878-419

Query Match      100.0%; Score 53; DB 16; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144

RESULT 10
US-10-723-860-2448
; Sequence 2448, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10723.860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2448
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-2448

Query Match      100.0%; Score 53; DB 16; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144

RESULT 11
US-10-751-736-67
; Sequence 67, Application US/10751736
```

```
; Publication No. US20040265230A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Martinez, Robert
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING COLON
; TITLE OF INVENTION: CANCERS
; FILE REFERENCE: AM100927 (031896-002000)
; CURRENT APPLICATION NUMBER: US/10751,736
; CURRENT FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000
; PRIOR FILING DATE: 2003-01-06
; NUMBER OF SEQ ID NOS: 54873
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-751-736-67

Query Match      100.0%; Score 53; DB 16; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144

RESULT 12
US-10-732-923-1435
; Sequence 1435, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 1435
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-732-923-1435

Query Match      100.0%; Score 53; DB 17; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
Db      134 LLIDDKGTIKL 144

RESULT 13
US-10-764-425-123
; Sequence 123, Application US/10764425
; Publication No. US20040146921A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; APPLICANT: Taylor, Ian
; TITLE OF INVENTION: EXPRESSION PROFILES FOR COLON CANCER AND METHODS OF USE
; FILE REFERENCE: 5151
; CURRENT APPLICATION NUMBER: US/10764,425
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: 60/442,582
; PRIOR FILING DATE: 2003-01-24
```



Search completed: November 18, 2005, 10:44:44  
Job time : 57.172 secs

```
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 123
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-425-123

Query Match      100.0%; Score 53; DB 18; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
        |||||
Db      134 LLIDDKGTIKL 144

RESULT 14
US-10-756-149-5363
; Sequence 5363, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756.149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5363
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-5363

Query Match      100.0%; Score 53; DB 18; Length 297;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
        |||||
Db      134 LLIDDKGTIKL 144

RESULT 15
US-10-732-923-1427
; Sequence 1427, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732.923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 1427
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Gallus gallus
US-10-732-923-1427

Query Match      90.6%; Score 48; DB 17; Length 303;
Best Local Similarity 90.9%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LLIDDKGTIKL 11
        |||||
Db      134 LLIDDKGVIKL 144
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:03:06 : Search time 48.7742 Seconds  
(without alignments)  
71.366 Million cell updates/sec

Title: US-10-006-177-10

Perfect score: 46

Sequence: 1 KIQELNYNL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_16Dec04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	9	5	ABG32311 Immunogen
2	46	100.0	9	5	ABG32310 HLA-A2 as
3	46	100.0	9	6	ABU03358 Human exp
4	46	100.0	171	4	AAB19969 Human STA
5	46	100.0	171	6	ABU04889 Human exp
6	46	100.0	268	4	AA772844 Mouse Sta
7	46	100.0	582	2	AAW62996 Human tru
8	46	100.0	582	6	ABU04748 Human exp
9	46	100.0	582	8	ADH57036 Truncated
10	46	100.0	680	6	ABR59713 Human sig
11	46	100.0	701	2	AA41335 84 kD ISG
12	46	100.0	701	6	ABU04741 Human exp
13	46	100.0	712	2	AA72079 Human Sta
14	46	100.0	712	2	AAW03170 Human Sta
15	46	100.0	712	2	AAW62995 Human Sta
16	46	100.0	712	6	ABU04747 Human exp
17	46	100.0	712	6	ABU04735 Human exp
18	46	100.0	712	6	ABU04745 Human exp
19	46	100.0	712	6	ABU04743 Human exp
20	46	100.0	712	8	ADH57035 Truncated
21	46	100.0	712	8	ADJ32352 Human STA
22	46	100.0	712	8	ADN04464 Antipsori
23	46	100.0	729	5	AAU84356 Protein S
24	46	100.0	739	2	AA41334 91 kD ISG
25	46	100.0	739	6	ABU04740 Human exp

26	46	100.0	749	5	AA78526 Rat STAT-
27	46	100.0	750	2	AA72078 Human Sta
28	46	100.0	750	2	AAW03168 Human STA
29	46	100.0	750	2	AAW62994 Human Sta
30	46	100.0	750	4	AAB19962 Human sig
31	46	100.0	750	5	AAE15172 Human STA
32	46	100.0	750	5	AA78525 Human STA
33	46	100.0	750	6	ABP98871 Human STA
34	46	100.0	750	6	ABU04749 Human exp
35	46	100.0	750	6	ABU04742 Human exp
36	46	100.0	750	6	ABU04744 Human exp
37	46	100.0	750	6	ABU04737 Human exp
38	46	100.0	750	6	ABU04736 Human exp
39	46	100.0	750	6	ABU04750 Human exp
40	46	100.0	750	6	ABU04746 Human exp
41	46	100.0	750	7	ADH57034 Human tra
42	46	100.0	750	8	ADH57034 Human STA
43	46	100.0	750	8	ADJ32351 Marker ge
44	46	100.0	750	8	ADJ75406 Antipsori
45	46	100.0	750	8	ADN03678 Antipsori

#### ALIGNMENTS

##### RESULT 1

ABG32311  
ID ABG32311 standard; peptide; 9 AA.

XX AC ABG32311;

XX DT 05-NOV-2002 (first entry)

XX DE Immunogenic peptide from human signal transducer and activator protein.

XX KW Human; immunogen; epitope; HLA-A1; human leukocyte antigen; CTL;

XX KW cytotoxic Y lymphocyte; cytostatic; cancer; colorectal carcinoma;

XX KW ovarian carcinoma; lung carcinoma; prostate carcinoma; vaccine; tumour;

XX KW HLA-2; passive immunotherapy; signal transducer and activator protein.

XX OS Homo sapiens.

XX PN WO200246416-A2.

XX PD 13-JUN-2002.

XX PF 04-DEC-2001; 2001WO-US047290.

XX PR 04-DEC-2000; 2000US-0251022P.

XX PR 20-DEC-2000; 2000US-0256824P.

XX PA (ARGO-) ARGONEX INC.

XX PI Ramakrishna V, Ross M, Philip R;

XX XX WPI; 2002-619021/66.

XX DR New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte,

XX PT and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.

XX PS Claim 1; Page 50; 60pp; English.

XX CC The invention relates to an immunogen comprising an isolated polypeptide

XX CC whose amino acid sequence comprises an epitopic peptide, does not include

XX CC MAGE 4 or MFG-E8 proteins, or consists of MAGE D protein or its

XX CC immunologically active fragment. Also included are a polynucleotide

XX CC encoding the immunogen or its complement, a vector comprising the

XX CC polynucleotide, a mammalian cell comprising the vector and expressing the

XX CC polynucleotide, a vaccine composition comprising the immunogen and an

XX CC antibody specific for the immunogen. The immunogen is useful for inducing

XX CC a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour

XX CC cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is

XX CC useful for inducing a CTL response when administered to a subject. A



PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 PS Claim 10; SEQ ID NO 138; 134pp; English.  
 XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 9A;

Query Match 100.0%; Score 46; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
 Db 1 KLOELNYNL 9  
 |||||

RESULT 4  
 AAB19969  
 ID AAB19969 standard; protein; 171 AA.

AC AAB19969;

DT 28-MAR-2001 (first entry)

XX Human STAT-1 DNA binding domain.

DE STAT-1; signal transducer and activator of transcription 1; human;  
 KW crystal; drug screening; DNA binding domain.

XX Homo sapiens.

XX US6160092-A.

XX 12-DEC-2000.

XX 29-MAY-1998; 98US-00087465.

XX 29-MAY-1998; 98US-00087465.

XX (UYRQ) UNIV ROCKEFELLER.

XX Chen X, Darnell JE, Kuriyan J, Vinkemeier U, Zhao Y, Jeruzalmi D;  
 PI WPI; 2001-101568/11.

DR N-PSDB; AAA89235.

XX Novel crystal useful in drug screening assays, comprises portion of  
 PT signal transducer, activator of transcription and duplex DNA.

PS Claim 9; Col 93-96; 206pp; English.

XX The present sequence is that of the DNA binding domain of human signal  
 CC transducer and activator of transcription 1 (STAT-1), i.e. amino acids  
 CC 318-488 of the full-length protein (see AAB19962). The domain includes an

CC immunoglobulin-type fold. The invention provides a crystal of a core  
 CC portion of a STAT protein in dimer form with an 18-mer duplex DNA (see  
 CC AAA89233) that contains a binding site for the STAT dimer. The core  
 CC portion comprises a coiled-coil domain, the DNA binding domain, a SH2  
 CC domain and a linker domain that joins the DNA binding and SH2 domains.  
 CC The crystal is of sufficient quality to perform X-ray crystallography  
 CC studies. Methods of preparing the crystals are included in the invention.  
 CC Knowledge of the STAT protein's 3-dimensional structure will aid in  
 CC structure-based drug design. The crystal can be used in drug screening  
 CC assays to identify agonist and antagonist compounds. Antagonists can be  
 CC used to treat inflammation, allergy, asthma and leukaemia, and agonists  
 CC to treat anaemia, neutropenia, thrombocytopenia, cancer, obesity, viral  
 CC diseases, growth retardation, and other conditions characterized by  
 CC insufficient STAT activity. Fusion proteins comprising a portion of STAT,  
 CC especially the DNA binding domain, and a fusion partner are also  
 CC disclosed

XX  
 SQ Sequence 171 AA;

Query Match 100.0%; Score 46; DB 4; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
 Db 34 KLOELNYNL 42  
 |||||

RESULT 5  
 ABU04889  
 ID ABU04889 standard; protein; 171 AA.

XX AC ABU04889;

XX 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1555.

DE Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCO INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.

XX Example 2; SEQ ID NO 1555; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a

CC fragment of a kinase, phosphatase, protease, protease inhibitor.  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 171 AA;

Query Match 100.0%; Score 46; DB 6; Length 171;  
Best Local Similarity 100.0%; Pred. No. 1.9; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 1 KQELNYNL 9  
| | | | |  
Db 34 KQELNYNL 42

RESULT 6  
AAV72844  
ID AAV72844 standard; protein; 268 AA.

AC AAV72844;

DT 31-MAY-2001 (first entry)

DE Mouse Stat1 protein fragment #2 (107-374 amino acids).

XX Mouse; Stat1 protein; transcription factor; c-Jun; gene transcription;  
KW cellular transformation; dysproliferative disease; cancer; psoriasis;  
KW therapy.

XX Mus musculus.

XX WO20011605-A2.

XX 08-MAR-2001.

XX 30-AUG-2000; 2000WO-US023822.

XX 31-AUG-1999; 99US-00387418.

XX (UVRQ ) UNIV ROCKEFELLER.

PI Zhang X, Horvath C, Wzesczynska MH, Darnell JE;

XX WPI; 2001-226705/23.

XX Identifying an agent for use in modulating the interaction between  
PT transcription factor c-Jun and a Stat3 protein.

PS Claim 65; Page 70-71; 86pp; English.

XX The present sequence is mouse Stat1 protein fragment containing 107-374  
CC amino acids of Stat1 protein. The invention relates to methods for  
CC identifying interacting regions of transcription factors and methods for  
CC identifying agents which modulates the interaction between a  
CC transcription factor such as c-Jun and a Stat protein such as Stat-1 and  
CC Stat-3, useful for modulating gene transcription e.g., cellular  
CC transformation. These identifying agents are used in the treatment of  
CC dysproliferative diseases and also for treating cancer and psoriasis. A  
CC Stat protein comprises the N-terminal domain, coiled-coil domain, DNA  
CC binding domain, linker domain, SH2 domain and transactivation domain

XX SQ Sequence 268 AA;

Query Match 100.0%; Score 46; DB 4; Length 268;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9  
| | | | |  
Db 244 KQELNYNL 252

RESULT 7

AAW62996

ID AAW62996 standard; protein; 582 AA.

XX AC AAW62996;

XX 24-NOV-1998 (first entry)

DT Human truncated Stat1 protein Stat1tc.

XX Stat1 protein; signal transducer and activator of transcription; human;  
DE purification; inflammation; allergy; asthma; leukaemia; anaemia;  
KW neutropenia; thrombocytopenia; cancer; obesity; growth retardation;  
KW drug screening; therapy; Stat1tc.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 24

XX Modified-site 309 /note= "optionally alkylated"

XX Modified-site 361 /note= "optionally alkylated"

XX Modified-site 570 /note= "optionally alkylated"

XX Modified-site 570 /label= O-phosphorylated

XX CA2218456-A.

XX 15-APR-1998.

XX 15-OCT-1997; 97CA-02218456.

XX 15-OCT-1996; 96US-0028176P.

XX (UVRQ ) UNIV ROCKEFELLER.

XX Vinkemeier U, Darnell JE;

XX WPI; 1998-399711/35.

XX N-PSDB; AAV42537.

XX New Stat protein and N-terminal fragment used in drug screening - to  
PT obtain products for treating e.g. inflammation, allergy, asthma, anaemia,  
PT cancer, obesity, viral disease and growth retardation.

XX Claim 1; Page 82-84; 111pp; English.

XX This is the amino acid sequence of a soluble, truncated, protease-  
CC resistant Stat1 protein, designated Stat1tc, that comprises amino acid  
CC residues 132-713 of human Stat1-alpha protein (see AAW62994). It does not  
CC include the N-terminal compact domain of the full-length protein but  
CC retains the Tyr701 phosphorylation site. The invention describes methods  
CC of producing milligram quantities of 3 forms of purified Stat1 protein:  
CC Stat1-alpha and Stat1-beta (see AAW62995) from baculovirus-infected  
CC insect cells, and Stat1tc from E. coli. In vitro phosphorylation of the  
CC Stat proteins plus cysteine residue alkylation to prevent aggregation,  
CC coupled to a chromatography protocol, has allowed the purification of  
CC activated Stat proteins. The Stat proteins and fragments can be used to  
CC identify antagonists and agonists of Stat function. Antagonists can be  
CC used to treat e.g. inflammation, allergy, asthma and leukaemias, and

CC agonists can be used in the treatment of e.g. anaemias, neutropaenias,  
CC thrombocytopenia, cancer, obesity, viral diseases, growth retardation or  
CC other diseases characterised by insufficient Stat activity  
XX  
SQ Sequence 582 AA;

Query Match 100.0%; Score 46; DB 2; Length 582;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLQELNYNL 9  
|||||  
Db 219 KLQELNYNL 227

RESULT 8  
ABU04748  
ID ABU04748 standard; protein; 582 AA.  
XX  
AC ABU04748;

DT 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1414.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW process; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

OS  
XX  
XX WO200278524-A2.

XX  
XX PD 10-OCT-2002.

XX  
XX PF 28-MAR-2002; 2002WO-US009671.

XX  
XX PR 28-MAR-2001; 2001US-0279495P.

XX  
XX PR 21-MAY-2001; 2001US-0292544P.

XX  
XX PR 08-AUG-2001; 2001US-0310801P.

XX  
XX PR 01-OCT-2001; 2001US-0326370P.

XX  
XX PR 04-DEC-2001; 2001US-0338780P.

XX  
XX PR 20-FEB-2002; 2002US-0358985P.

XX  
XX PA (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX  
XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.

XX Example 2; SEQ ID NO 1414; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this  
XX polypeptide, is useful for treating cancer. The polypeptide is also  
XX useful for identifying compounds that binds to a naturally processed  
XX class I or class II MHC-binding polypeptide. The polypeptides and  
XX polynucleotides are particularly useful for treating or preventing  
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
XX lymphoma or leukaemia. These are also useful for screening agents for  
XX treating the above mentioned diseases. This sequence represents an  
XX expressed protein tag (EPT) isolated from human tissue for translational

CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 582 AA;

Query Match 100.0%; Score 46; DB 6; Length 582;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLQELNYNL 9  
|||||

Db 219 KLQELNYNL 227

RESULT 9  
ADH57036

ID ADH57036 standard; protein; 582 AA.

XX  
AC ADH57036;

XX  
XX DT 25-MAR-2004 (first entry)

XX Truncated human Stat1tc protein SeqID 3.

XX human; Stat1tc; Stat; transcription factor;  
KW signal transducers and activators of transcription; cytokine;  
KW growth factor; protein aggregation; drug discovery; inclusion body.

XX Homo sapiens.

XX  
XX PN US2003092066-A1.

XX  
XX PD 15-MAY-2003.

XX  
XX PF 17-SEP-2002; 2002US-00245120.

XX  
XX PR 02-NOV-1999; 99US-00430806.

XX  
XX PA (VINK/) VINKEMEIER U.

XX  
XX PA (DARN/) DARNELL J E.

XX  
XX PI Vinkemeier U, Darnell JE;

XX  
XX DR WPI; 2004-068840/07.

XX  
XX DR N-PSDB; ADH57038.

XX Novel truncated signal transducers and activators of transcription  
XX protein and a purified N-terminal peptide fragment of the protein useful  
XX for drug discovery.

XX Claim 1; SEQ ID NO 3; 55pp; English.

XX This invention relates to novel methods of purifying recombinant Stat  
XX proteins, as well as modified and functional fragments thereof.  
XX Specifically, it refers to the protein family of transcription factors  
XX identified as the signal transducers and activators of transcription  
XX (Stat) that mediate the action of signalling molecules including  
XX cytokines and growth factors. The present invention describes various  
XX methods for producing milligram quantities of different forms of purified  
XX Stat proteins from recombinant DNA constructs, which are produced in  
XX insect cells infected with recombinant baculovirus. As such, it provides  
XX for the isolation of active tyrosine phosphorylated Stat proteins, as  
XX well as unphosphorylated Stat, truncated Stat proteins and those that  
XX exhibit modifications of specific cysteine residues that prevents protein  
XX aggregation. These recombinant Stat proteins are useful in the drug  
XX discovery process and this method that expresses the truncated protein in  
XX a soluble form overcomes earlier failures whereby recombinant Stat  
XX proteins were almost exclusively formed as insoluble inclusion bodies.  
XX This polypeptide sequence is the truncated human Stat1tc protein of the  
XX invention.

XX  
SQ Sequence 582 AA;

Query Match 100.0%; Score 46; DB 8; Length 582;  
Best Local Similarity 100.0%; Pred. No. 6.6; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 1 KQELNYNL 9  
| | | | | | | |  
Db 219 KQELNYNL 227

RESULT 10  
ABR59713  
ID ABR59713 standard; protein; 680 AA.  
XX  
AC ABR59713;  
XX  
DT 25-JUL-2003 (first entry)  
XX  
DE Human signal transducer and activator of transcription 1 isoform alpha.  
XX  
DE Human; T lymphocyte activation; T-cell; A-raf-1; TCPTP/PTPN2; asthma;  
KW immunosuppressive; antiasthmatic; antiallergic; antiinflammatory;  
KW lymphocyte activation; lymphocyte migration; cytokine production;  
KW cell surface marker expression; antibody production; apoptosis; allergy;  
KW antibody proliferation; antibody differentiation; hypersensitivity;  
KW graft versus host disease; inflammation; transducer;  
KW transcriptional activator.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029277-A2.  
XX  
PD 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US031618.  
XX  
PR 03-OCT-2001; 2001US-0327212P.  
XX  
PA (RIGE-) RIGEL PHARM INC.  
XX  
PI Chu P, Li C, Liao XC, Masuda E, Pardo J, Zhao H;  
XX WPI; 2003-363276/34.  
XX N-PSDB; ACC81108.  
XX  
PT Identifying a compound that modulates T lymphocyte activation, useful for  
PT monitoring changes in cell surface marker expression, comprises  
PT contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with  
PT a compound.  
XX  
PS Disclosure; Page 88; 126pp; English.  
XX

The invention relates to a novel method for identifying a compound that modulates T lymphocyte activation. The method comprises contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with a compound, where the A-raf-1 or TCPTP/PTPN2 polypeptide is encoded by a nucleic acid that hybridizes to a nucleic acid encoding a polypeptide having a sequence selected from two 606-amino acid sequence and a 415-amino acid sequence given in the specification. The method of the invention has immunosuppressive, antiasthmatic, antiallergic, and antiinflammatory activity. The method is useful for identifying compounds that modulate lymphocyte activation and migration, and for monitoring changes in cell surface marker expression, cytokine production, antibody production, cell proliferation and differentiation, and apoptosis, using either cell lines or primary cells. The A-raf-1 or TCPTP/PTPN2 proteins may be used as drug targets for compounds that suppress or activate lymphocyte activation and migration, e.g. for the treatment of diseases in which modulation of the immune response is desired such as delayed type hypersensitivity reactions, asthma, allergies, graft versus host disease, and acute and chronic inflammation. Modulators of lymphocyte activation are useful for treating disorders related T and B cell activation and migration. The present sequence is used in the exemplification of the invention

XX SQ Sequence 680 AA;  
Query Match 100.0%; Score 46; DB 6; Length 680;  
Best Local Similarity 100.0%; Pred. No. 7.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9  
| | | | | | | |  
Db 350 KQELNYNL 358

RESULT 11  
AAR41335  
ID AAR41335 standard; protein; 701 AA.  
XX  
AC AAR41335;  
XX  
DT 25-MAR-2003 (revised)  
DT 22-APR-1994 (first entry)  
XX  
DE 84 kD ISGF-3alpha.  
XX  
DE 113 kD; 91 kD; 84 kD; ISGF-3alpha; interferon-related; receptor;  
KW recognition factor; gene family; translation protein; tyrosine;  
KW DNA binding protein; interferon-gamma; hairy cell leukaemia;  
KW interferon therapy; chronic viral hepatitis; phosphorylation;  
KW adjuvant therapy; tyrosine kinase.  
XX  
OS Homo sapiens.  
XX  
PN WO9319179-A1.  
XX  
PD 30-SEP-1993.  
XX  
PF 19-MAR-1993; 93WO-US002569.  
XX  
PR 19-MAR-1992; 92US-00854296.  
PR 23-NOV-1992; 92US-00980498.  
XX  
PA (UYRQ) UNIV ROCKEFELLER.  
XX  
PI Darnell JE, Schindler CW, Fu X, Shuai K;  
XX WPI; 1993-320745/40.  
XX N-PSDB; AAQ49166.  
XX  
PT Interferon receptor recognition factors - useful e.g. to treat viral  
PT hepatitis, hairy cell leukaemia and to potentiate interferon effects.  
XX  
PS Claim 17; Fig 3; 131pp; English.  
XX

The sequences given in AAR41333-35 represent the 113 kD, 91 kD and 84 kD ISGF-3alpha proteins respectively. ISGF-3alpha is an interferon-related receptor recognition factor which comprises several substituents. The 113 kD, and the 91 and 84 kD proteins are derived from two different but related genes. It is clear that a gene family exists and further members are likely to be found. The 91 kD protein has the capability of acting as a translation protein and as a DNA binding protein in response to interferon-gamma stimulation. These proteins participate in rapid phosphorylation and dephosphorylation during the course of, and as part of their activity. This phosphorylation takes place in an interferon-dependant manner on specified tyrosine residues. These proteins may be used in conjunction with interferon therapy eg. to treat chronic viral hepatitis, hairy cell leukaemia and for use with interferon in adjuvant therapy. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 701 AA;  
Query Match 100.0%; Score 46; DB 2; Length 701;  
Best Local Similarity 100.0%; Pred. No. 8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
Qy      1 KLOELNYNL 9
Db      339 KLOELNYNL 347

RESULT 12
ABU04741
ID ABU04741 standard; protein; 701 AA.
XX
AC ABU04741;
XX
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1407.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
XX WO200278524-A2.
XX
XX 10-OCT-2002.
XX
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
XX
XX 21-MAY-2001; 2001US-0292544P.
XX
XX 08-AUG-2001; 2001US-0310801P.
XX
XX 01-OCT-2001; 2001US-0326370P.
XX
XX 04-DEC-2001; 2001US-0336780P.
XX
XX 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCO INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1407; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 701 AA;

Query Match      100.0%; Score 46; DB 6; Length 701;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      339 KLOELNYNL 347

RESULT 13
AAR72079
ID AAR72079 standard; protein; 712 AA.
XX
AC AAR72079;
XX
XX
DT 25-MAR-2003 (revised)
DT 27-SEP-1995 (first entry)
XX
DE Human Stat84.
XX
KW Signal transducer and activator of transcription; ISGF-3; STAT; Stat84;
KW receptor recognition factor; transcription factor; cellular debilitation;
KW derangement; dysfunction; interferon-gamma.
XX
XX Homo sapiens.
XX
XX WO9508629-A1.
XX
XX 30-MAR-1995.
XX
XX 26-SEP-1994; 94WO-US010849.
XX
XX 24-SEP-1993; 93US-00126588.
XX
XX 24-SEP-1993; 93US-00126595.
XX
XX 11-MAR-1994; 94US-00212184.
XX
XX 11-MAR-1994; 94US-00212185.
XX
XX (UYRQ) UNIV ROCKEFELLER.
XX
XX Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;
XX
XX WPI; 1995-139598/18.
XX
XX N-PSDB; AAQ89337.
XX
XX Receptor recognition factor implicated in transcriptional stimulation of
XX genes - useful in drug screening assays and/or for treating cellular
XX debilitations, derangements and/or dysfunctions, etc.
XX
XX Disclosure; Page 91-94; 160pp; English.
XX
XX The sequences of cDNA encoding receptor recognition factors having mol.wt.
XX of 113 kDa (Stat13), 91 kDa (Stat92) and 84 kDa (Stat84) are given in
XX AAQ89335-37 and the deduced amino acid sequences of the STAT proteins in
XX AAR72077-79. These ISGF-3-derived proteins are activated by binding of
XX interferon-alpha (all 3 Stat proteins) or interferon-gamma (Stat91) to
XX cell receptors. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 712 AA;

Query Match      100.0%; Score 46; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 350

RESULT 14
AAW03170
ID AAW03170 standard; protein; 712 AA.
XX
AC AAW03170;
XX
XX
DT 24-OCT-1996 (first entry)
XX
DE Human STAT1-beta.
XX
```

```

XX STAT1-beta; signal transducer and activator of transcription;
KW DNA binding protein; ligand: receptor; oncogenesis; inflammation;
KW autoimmune disease; antagonist; therapy; STAT84.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 399..508
FT /label= DNA binding_domain
FT /note= "Claim 3, page 110"
XX
XX WO9620954-A2.
XX
XX 11-JUL-1996.
XX
XX 28-DEC-1995; 95WO-US017025.
XX
XX 06-JAN-1995; 95US-00369796.
XX
XX (UYRQ ) UNIV ROCKEFELLER.
XX
XX Darnell JE, Wen Z, Horvath CM, Zhong Z;
XX
XX WPI; 1996-333941/33.
XX
XX N-PSDB; AAT31277.
XX
XX New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,
XX preventing or treating cellular dysfunction, e.g. oncogenesis,
XX inflammation, parasitic disease or autoimmunity.
XX
XX Disclosure; Page 76-79; 138pp; English.
XX
XX Signal transducer and activator of transcription (STAT) protein STAT1-
XX beta (AAW03170), also known as STAT84, is a 84 kDa protein having a dual
XX purpose, i.e. signal transduction from ligand-activated receptor kinase
XX complexes followed by nuclear translocation and DNA binding to activate
XX transcription. Recombinant STAT1-beta can be obtained using an isolated cDNA
XX clone (AAT31277). STAT1-beta includes a DNA-binding domain (see also
XX AAW03165) capable of both receptor recognition and message delivery via
XX DNA binding in a receptor-ligand specific manner. STAT1-beta is a
XX truncated form of STAT1-alpha (AAW03168). STAT proteins and their DNA
XX binding domains (see also AAW03165-69, AAW03171-76) are useful for
XX screening antagonists used to inhibit STAT-mediated signal transduction
XX and activation of transcription
XX
XX Sequence 712 AA;
XX
XX Query Match 100.0%; Score 46; DB 2; Length 712;
XX Best Local Similarity 100.0%; Pred. No. 8.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 KLOELNYNL 9
XX |||||
XX 350 KLOELNYNL 358
XX
XX RESULT 15
XX AAW62995
XX ID AAW62995 standard; protein; 712 AA.
XX
XX AC AAW62995;
XX
XX DT 24-NOV-1998 (first entry)
XX
XX DE Human Stat1-beta protein.
XX
XX KW Stat1 protein; signal transducer and activator of transcription; human;
KW purification; inflammation; allergy; asthma; leukaemia; anaemia;
KW neutropenia; thrombocytopenia; cancer; obesity; growth retardation;
KW drug screening; therapy.
XX
XX OS Homo sapiens.

```

```

XX Key Location/Qualifiers
FH Domain 1..131
FT /label= Compact_domain
FT Modified-site 155
FT /note= "optionally alkylated"
FT Modified-site 440
FT /note= "optionally alkylated"
FT Modified-site 492
FT /note= "optionally alkylated"
FT Modified-site 701
FT /label= O-phosphorylated
XX
XX CA2218456-A.
XX
XX 15-APR-1998.
XX
XX 15-OCT-1997; 97CA-02218456.
XX
XX 15-OCT-1996; 96US-0028176P.
XX
XX (UYRQ ) UNIV ROCKEFELLER.
XX
XX Vinkemeier U, Darnell JE;
XX
XX WPI; 1998-399711/35.
XX
XX New Stat protein and N-terminal fragment used in drug screening - to
XX obtain products for treating e.g. inflammation, allergy, asthma, anaemia,
XX cancer, obesity, viral disease and growth retardation.
XX
XX Claim 20; Page 80-82; 111pp; English.
XX
XX This is human Stat1-beta protein, a major proteolytic cleavage fragment
XX (aa1-712) of Stat1-alpha protein (see AAW62994). Stat1-beta is resistant
XX to further proteolysis. The invention describes methods of producing
XX milligram quantities of 3 forms of purified Stat1 protein: Stat1-alpha
XX and Stat1-beta from baculovirus-infected insect cells, and truncated
XX Stat1c (see AAW62996) from E. coli. Stat1-beta has an N-terminal compact
XX domain that enhances the DNA binding of the Stat protein. The Tyr-701
XX residue is phosphorylated in vivo. In vitro phosphorylation of the
XX protein plus cysteine residue alkylation to prevent aggregation, coupled
XX to a chromatography protocol, has allowed the purification of activated
XX Stat proteins. The Stat proteins and fragments can be used to identify
XX antagonists and agonists of Stat function. Antagonists can be used to
XX treat e.g. inflammation, allergy, asthma and leukaemias, and agonists can
XX be used in the treatment of e.g. anaemias, neutropenias,
XX thrombocytopenia, cancer, obesity, viral diseases, growth retardation or
XX other diseases characterised by insufficient Stat activity
XX
XX Sequence 712 AA;
XX
XX Query Match 100.0%; Score 46; DB 2; Length 712;
XX Best Local Similarity 100.0%; Pred. No. 8.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 KLOELNYNL 9
XX |||||
XX 350 KLOELNYNL 358
XX
XX Search completed: November 18, 2005, 01:01:28
XX Job time : 51.7742 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:11:06 ; Search time 8.90323 Seconds  
(without alignments)  
97.263 Million cell updates/sec

Title: US-10-006-177-10

Perfect score: 46

Sequence: 1 KQELNYNL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	46	100.0	739	2 A46159	interferon-depende
2	38	82.6	354	2 B97120	DNA uptake protein
3	36	78.3	345	2 T32203	hypothetical prote
4	36	78.3	487	2 B95059	hypothetical prote
5	36	78.3	487	2 A97928	type I site-specif
6	36	78.3	674	2 T28274	ORF MSV113 probabl
7	35	76.1	69	2 B83982	hypothetical prote
8	35	76.1	129	2 S41972	14.9K protein - fo
9	35	76.1	333	2 A95039	sugar binding tran
10	35	76.1	355	2 B97909	transcription regu
11	35	76.1	478	2 T33735	hypothetical prote
12	35	76.1	1005	2 A64455	hypothetical prote
13	34	73.9	120	2 F81802	probable transcrip
14	34	73.9	143	2 C81065	transcription regu
15	34	73.9	417	2 G98200	sarcosine oxidase
16	34	73.9	417	2 AB3086	sarcosine oxidase
17	34	73.9	432	2 S26432	intermediate filam
18	34	73.9	567	2 T18872	intermediate filam
19	34	73.9	618	2 A80770	probable outer mem
20	34	73.9	1008	2 AE2304	hypothetical prote
21	34	73.9	2573	2 D71614	hypothetical prote
22	33	71.7	180	2 C71869	hypothetical prote
23	33	71.7	216	2 AC1707	L-fuculose-phospha
24	33	71.7	216	2 AF1336	L-fuculose-phospha
25	33	71.7	224	2 B91111	hypothetical prote
26	33	71.7	224	2 E85956	hypothetical prote
27	33	71.7	532	2 T01418	pectinesterase hom
28	33	71.7	1165	2 A46180	adenyl cyclase t
29	33	71.7	1165	2 A49201	adenyl cyclase t

30	33	71.7	1171	2 T17454	diaphanous-related
31	32	69.6	133	2 T30474	hypothetical prote
32	32	69.6	154	2 S36993	transposase (clone
33	32	69.6	159	2 A97836	hypothetical prote
34	32	69.6	232	2 F72383	hypothetical prote
35	32	69.6	250	2 S44768	C2984.7 protein -
36	32	69.6	261	2 G90449	conserved hypothet
37	32	69.6	271	2 H72261	conserved hypothet
38	32	69.6	336	2 S04682	ribosomal protein
39	32	69.6	357	2 S58215	jasmonate induced
40	32	69.6	367	2 A99468	oxidoreductase (im
41	32	69.6	375	2 D97268	toxic anion resist
42	32	69.6	420	2 S19707	site-specific DNA-
43	32	69.6	486	2 A75045	2-isopropylmalate
44	32	69.6	507	2 AG3187	aldehyde dehydroge
45	32	69.6	528	2 E71886	type I restriction

ALIGNMENTS

RESULT 1

A46159

interferon-dependent positive-acting transcription factor ISGF-3 91K chain - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A46159

R:Schindler, C.; Fu, X.Y.; Improt, T.; Aebersold, R.; Darnell Jr., J.E.  
Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992  
A:Title: Proteins of transcription factor ISGF-3: one gene encodes the 91-and 84-kDa ISG

A:Reference number: A46159; MUID:92366557; PMID:1502203

A:Accession: A46159

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid; protein

A:Residues: 1-739 <SCH>

A:Cross-references: UNIPROT:P42224

A:Experimental source: HeLa cells

A:Note: sequence extracted from NCBI backbone (NCBIP:110818)

C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 100.0%; Score 46; DB 2; Length 739;  
Best Local Similarity 100.0%; Pred. No. 0.58;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KQELNYNL 9  
|||  
Db 339 KQELNYNL 347

RESULT 2

B97120

DNA uptake protein [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004

C:Accession: B97120

R:Nolling, J.; Bretton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: B97120

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-354 <KUR>

A:Cross-references: UNIPROT:Q97169; GB:AE001437; PID:AAK79749.1; PID:GI5024755; GSPDB:G

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC1784

Query Match 82.6%; Score 38; DB 2; Length 354;  
Best Local Similarity 77.8%; Pred. No. 10;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 KQLELNYNL 9
|| |||||
Db 95 KLELNYNV 103

RESULT 3
T32203
hypothetical protein T02B11.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T32203
R/Goela, D.
submitted to the EMBL Data Library, September 1997
A/Description: The sequence of C. elegans cosmid T02B11.
A/Reference number: Z21135
A/Accession: T32203
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-345 <GB>
A/Cross-references: UNIPROT:O16975; EMBL:AF022979; PIDN:AA669906.1; GSPDB:GN000023; CESP:
A/Experimental source: strain Bristol N2; clone T02B11
C/Genetics:
A/Gene: CESP:T02B11.5
A/Map position: 5
A/Introns: 68/3; 197/2; 237/3; 294/2; 343/3

Query Match 78.3%; Score 36; DB 2; Length 345;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQLELNYNL 9
|| |||||
Db 217 KQKINFNL 225

RESULT 4
B95059
hypothetical protein SP0509 [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C/Accession: B95059
R/Tetcelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: B95059
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-487 <KUR>
A/Cross-references: UNIPROT:Q97879; GB:AE005672; PIDN:AAK74667.1; PID:g14971982; GSPDB:G
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SP0509
C/Superfamily: type I site-specific deoxyribonuclease chain hsdM

Query Match 78.3%; Score 36; DB 2; Length 487;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KQLELNYNL 9
|| |||||
Db 432 ELAELNYNL 440

RESULT 5
A97928
type I site-specific deoxyribonuclease (EC 3.1.21.3) chain hsdM [imported] - Streptococ
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: A97928
```

```
R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; MUID:21429245; PMID:11544234
A/Accession: A97928
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-487 <KUR>
A/Cross-references: UNIPROT:Q8DQX1; GB:AE007317; PIDN:AAK99253.1; PID:g15458017; GSPDB:B
A/Genes: hsdM
C/Superfamily: type I site-specific deoxyribonuclease chain hsdM

Query Match 78.3%; Score 36; DB 2; Length 487;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KQLELNYNL 9
|| |||||
Db 432 ELAELNYNL 440

RESULT 6
T28274
ORF MSV113 probable early transcription factor small subunit VETP-S homolog (vaccinia D6
C/Species: Melanoplus sanguinipes entomopoxvirus
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C/Accession: T28274
R/Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A/Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A/Reference number: Z20484; MUID:99102612; PMID:9847359
A/Accession: T28274
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-674 <AFO>
A/Cross-references: UNIPROT:Q9VYX9; EMBL:AF063866; NID:g4049647; PIDN:AAK97658.1; PID:g4
C/Genetics:
A/Note: MSV113
C/Superfamily: vaccinia virus early transcription factor 70K chain

Query Match 78.3%; Score 36; DB 2; Length 674;
Best Local Similarity 77.8%; Pred. No. 48;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KQLELNYNL 9
|| |||||
Db 262 KQEDNYNM 270

RESULT 7
B83982
hypothetical protein BH2661 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: B83982
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: B83982
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-69 <STO>
A/Cross-references: UNIPROT:Q9K916; GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BA063
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH2661

Query Match 76.1%; Score 35; DB 2; Length 69;
```

Best Local Similarity 77.8%; Pred. No. 7.1; Mismatches 0; Indels 2; Gaps 0;

Query 1 KLQELNYNL 9  
 Db 36 KLSKKNYNL 44

RESULT 8  
 S41972  
 14.9K protein - fowlpox virus  
 C:Species: fowlpox virus  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
 C:Accession: S41972  
 R:Skinner, M.A.; Moore, J.B.; Binns, M.M.; Boursnell, M.E.  
 submitted to the EMBL Data Library, February 1994  
 A:Description: Deletion of fowlpox virus homologues of vaccinia virus genes between the  
 A:Reference number: S41971  
 A:Accession: S41972  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-129 <SKI>  
 A:Cross-references: EMBL:Z29716; NID:G453598; PIDN:CAA82803.1; PID:G453600  
 C:Superfamily: fowlpox virus 14.9K protein

Query Match 76.1%; Score 35; DB 2; Length 129;  
 Best Local Similarity 75.0%; Pred. No. 14; Mismatches 1; Indels 0; Gaps 0;

Query 1 KLQELNYN 8  
 Db 90 KMSLNYN 97

RESULT 9  
 A95039  
 sugar binding transcription regulator RegR [imported] - Streptococcus pneumoniae (strain  
 C:Species: Streptococcus pneumoniae  
 C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
 C:Accession: A95039  
 R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
 on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf, E.  
 nson, T.; Hickey, E.K.; Holt, I.E.  
 Science 293, 498-506, 2001  
 A:Authors: Lofthus, B.J.; Yang, F.; Smith, J.C.; Dougherty, B.A.; Morrison,  
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
 A:Reference number: A95000; MUID:21357209; PMID:11463916  
 A:Accession: A95039  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-333 <KUR>  
 A:Cross-references: UNIPROT:Q97SK3; GB:AE005672; PIDN:AAK74506.1; PID:G14971804; GSPDB:G  
 A:Experimental source: strain TIGR4  
 C:Genetics:  
 A:Gene: SP0330

Query Match 76.1%; Score 35; DB 2; Length 333;  
 Best Local Similarity 75.0%; Pred. No. 36; Mismatches 2; Indels 0; Gaps 0;

Query 2 LOELNYNL 9  
 Db 260 IKELNYNL 267

RESULT 10  
 B97909  
 transcription regulator, member of GalR family regR [imported] - Streptococcus pneumoniae  
 C:Species: Streptococcus pneumoniae  
 C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
 C:Accession: B97909  
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

Y, P.; Sun, P.M.; Winkler, M.E.  
 J. Bacteriol. 183, 5709-5717, 2001  
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
 A:Reference number: A97872; MUID:21429245; PMID:11544234  
 A:Accession: B97909  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-355 <KUR>  
 A:Cross-references: UNIPROT:Q8DR72; GB:AE007317; PIDN:AAK99102.1; PID:G15457851; GSPDB:G  
 C:Genetics:  
 A:Gene: regR

Query Match 76.1%; Score 35; DB 2; Length 355;  
 Best Local Similarity 75.0%; Pred. No. 39; Mismatches 2; Indels 0; Gaps 0;

Query 2 LOELNYNL 9  
 Db 282 IKELNYNL 289

RESULT 11  
 T33735  
 hypothetical protein H34C03.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T33735  
 R:Kalicki, J.; Gibson, A.  
 submitted to the EMBL Data Library, October 1998  
 A:Description: The sequence of C. elegans cosmid H34C03.  
 A:Reference number: Z21394  
 A:Accession: T33735  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-478 <KAL>  
 A:Cross-references: UNIPROT:Q9TYV7; EMBL:AF100662; PIDN:AAC68976.1; GSPDB:GN00022; CESP  
 A:Experimental source: strain Bristol N2; clone H34C03  
 C:Genetics:  
 A:Gene: CESP:H34C03.1  
 A:Map position: 4  
 A:Introns: 57/3; 103/2; 185/3; 283/2; 347/3; 426/3  
 C:Superfamily: Caenorhabditis elegans hypothetical protein H34C03.1

Query Match 76.1%; Score 35; DB 2; Length 478;  
 Best Local Similarity 77.8%; Pred. No. 53; Mismatches 1; Indels 0; Gaps 0;

Query 1 KLQELNYNL 9  
 Db 226 RLQELMYNL 234

RESULT 12  
 A64465  
 hypothetical protein MJ1322 - Methanococcus jannaschii  
 C:Species: Methanococcus jannaschii  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
 C:Accession: A64465  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
 rson, J.D.; Sadow, P.W.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
 Reich, C.I.; Overbeek, R.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
 A:Reference number: A64300; MUID:96337999; PMID:8688087  
 A:Accession: A64465  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1005 <BUL>  
 A:Cross-references: UNIPROT:Q58718; GB:U67572; GB:L77117; NID:G1591958; PIDN:AA899331.1;  
 C:Genetics:  
 A:Map position: REV1273394-1270377

C;Superfamily: hypothetical protein M1322

Query Match 76.1%; Score 35; DB 2; Length 1005;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KQELNYN 8

Db 778 KLDEINYN 785

# RESULT 13

F81802 probable transcription regulator NMA1774 [imported] - Neisseria meningitidis (strain Z24)

C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004

C;Accession: F81802

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: F81802

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-120 <PAR>

A;Cross-references: UNIPROT:Q9JTH6; GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CA88500

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: NMA1774

Query Match 73.9%; Score 34; DB 2; Length 120;  
Best Local Similarity 85.7%; Pred. No. 20;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QELNYNL 9

Db 25 QDLNYNL 31

# RESULT 14

C81065 transacription regulator, MarR family NMB1585 [imported] - Neisseria meningitidis (strain

C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004

C;Accession: C81065

R;Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755; PMID:10710307

A;Accession: C81065

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-143 <TET>

A;Cross-references: UNIPROT:Q9JYH5; GB:AE002509; GB:AE002098; NID:g7226932; PIDN:AAF4193

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB1585

Query Match 73.9%; Score 34; DB 2; Length 143;  
Best Local Similarity 85.7%; Pred. No. 24;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QELNYNL 9

Db 25 QDLNYNL 31

# RESULT 15

G98200

sarcosine oxidase beta chain PA5416 [imported] - Agrobacterium tumefaciens (strain C58,  
C;Species: Agrobacterium tumefaciens

C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004

C;Accession: G98200

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markeiz, B.;

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: G98200

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-417 <KUR>

A;Cross-references: UNIPROT:Q8U7Y9; GB:AE007870; PIDN:AAK89129.1; PID:gl5158937; GSPDB:G

C;Genetics:

A;Gene: AGR\_L\_1106

A;Map position: linear chromosome

C;Superfamily: sarcosine oxidase

Query Match 73.9%; Score 34; DB 2; Length 417;

Best Local Similarity 85.7%; Pred. No. 72;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QELNYNL 9

Db 105 QELNYN 111

Search completed: November 18, 2005, 01:09:44

Job time : 12.0032 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:04:26 ; Search time 39.9677 Seconds  
(without alignments)  
115.311 Million cell updates/sec

Title: US-10-006-177-10

Perfect score: 46

Sequence: 1 KLOELNYNL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_crembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	712	Q6P6Q7	Q6P6Q7 rattus norv
2	46	100.0	712	Q99K94	Q99K94 mus musculus
3	46	100.0	749	Q8C3V4	Q8C3V4 mus musculus
4	46	100.0	749	Q8C497	Q8C497 mus musculus
5	46	100.0	749	Q9D323	Q9D323 mus musculus
6	46	100.0	749	Q9D323	Q9D323 mus musculus
7	46	100.0	750	Q9QXK0	Q9QXK0 rattus norv
8	46	100.0	750	STAL_HUMAN	P42224 homo sapien
9	46	100.0	755	Q88D00	Q88D00 homo sapien
10	46	100.0	1165	Q8C8M3	Q8C8M3 mus musculus
11	42	91.3	757	Q7TP57	Q7TP57 rattus norv
12	40	87.0	751	Q764W5	Q764W5 sus scrofa
13	38	82.6	354	Q8JGNO	Q8JGNO xenopus lae
14	36	78.3	115	Q8JMC3	Q8JMC3 mamestra co
15	36	78.3	117	Q7IAH8	Q7IAH8 mamestra co
16	36	78.3	117	Q7IAH8	Q7IAH8 mamestra co
17	36	78.3	154	Q9J8C0	Q9J8C0 spodoptera
18	36	78.3	175	Q8BI02	Q8BI02 mus musculus
19	36	78.3	197	Q8N124	Q8N124 homo sapien
20	36	78.3	269	Q8CFH8	Q8CFH8 mus musculus
21	36	78.3	277	Q87HQ0	Q87HQ0 vibrato para
22	36	78.3	285	Q8XN77	Q8XN77 clostroidium
23	36	78.3	399	Q8TBR2	Q8TBR2 homo sapien
24	36	78.3	441	Q8N8L3	Q8N8L3 homo sapien
25	36	78.3	487	Q8T879	Q8T879 streptococc
26	36	78.3	487	Q8DQX1	Q8DQX1 streptococc
27	36	78.3	538	Q95K16	Q95K16 macaca fasc
28	36	78.3	630	Q6CKZ4	Q6CKZ4 kluyveromyc
29	36	78.3	634	Q95JX7	Q95JX7 macaca fasc
30	36	78.3	674	ETFL_MSEPV	Q9YXN0 melanoplus
31	36	78.3	689	Q9FMH0	Q9FMH0 arabidopsis

32	36	78.3	737	2	Q6ZR13	Q6ZR13 homo sapien
33	36	78.3	754	2	O13131	O13131 oncorhynch
34	36	78.3	758	2	Q90Y17	Q90Y17 tetraodon f
35	36	78.3	764	2	Q90Y16	Q90Y16 tetraodon f
36	36	78.3	823	2	Q6SJRO	Q6SJRO arabidopsis
37	36	78.3	833	2	Q6P1W1	Q6P1W1 homo sapien
38	36	78.3	1326	2	Q6TFL3	Q6TFL3 homo sapien
39	35	76.1	69	2	Q9K9I6	Q9K9I6 bacillus ha
40	35	76.1	97	2	Q661H1	Q661H1 borrelia ga
41	35	76.1	120	2	Q9HKJ7	Q9HKJ7 thermoplasm
42	35	76.1	162	2	Q7LHX9	Q7LHX9 lactobacill
43	35	76.1	167	2	Q6G502	Q6G502 batonella
44	35	76.1	240	2	O82095	O82095 ceratopter
45	35	76.1	244	2	Q94Z17	Q94Z17 pyataiella 1

#### ALIGNMENTS

RESULT 1  
Q6P6Q7  
ID Q6P6Q7 PRELIMINARY; PRT; 712 AA.  
AC Q6P6Q7;  
DT 05-JUL-2004 (TRENBLrel. 27, Created)  
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)  
DE Stat1 protein.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T.I., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.B., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences".  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[2]  
SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RX Strausberg R.;  
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
EMBL; BC062079; AA062079.1; --  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0004871; F:signal transducer activity; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR008967; P53\_like\_DNA\_bnd.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001217; STAT.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF01017; STAT\_alpha; 1.  
DR Pfam; PF02864; STAT\_bind; 1.  
DR Pfam; PF02865; STAT\_int; 1.

```

DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 712 AA; 83110 MW; 193C7E1A1DBECBCC CRC64;

Query Match 100.0%; Score 46; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9
Db 350 KQELNYNL 358

RESULT 2
Q99K94
ID Q99K94 PRELIMINARY; PRT; 712 AA.
AC Q99K94;
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Stat1 protein (Mus musculus adult male spinal cord cDNA, RIKEN full-
DE length enriched library, clone:A30045H06 product:signal transducer
DE and activator of transcription 1, full insert sequence).
GN Name=Stat1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2ECH II; TISSUE=Mammary tumor;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C2ECH II; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44 (1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690 (2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;

```

```

RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
RN Nature 420:563-573 (2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes."
RN Genome Res. 10:1617-1630 (2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer."
RN Genome Res. 10:1757-1771 (2000).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hasegaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Satoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004808; AA04808.1; -
DR EMBL; AK039458; BA030355.1; -
DR HSSP; P42224; 1BF5.
DR MGD; MGI:103063; Stat1.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_inc; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 712 AA; 83106 MW; D62081709638CDA5 CRC64;

Query Match 100.0%; Score 46; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9
Db 350 KQELNYNL 358

RESULT 3
Q8C3V4
ID Q8C3V4 PRELIMINARY; PRT; 749 AA.

```



DR	HSSP; P42224; 1BF5.	
DR	MGD; MGI:103063; Stat1.	
DR	GO; GO:0005737; C:cytoplasm; IDA.	
DR	GO; GO:0005634; C:nucleus; IDA.	
DR	GO; GO:0005515; F:protein binding; IPI.	
DR	InterPro; IPR008967; P53_like_DNA_bnd.	
DR	InterPro; IPR000980; SH2.	
DR	InterPro; IPR001217; STAT.	
DR	Pfam; PF00017; SH2; 1.	
DR	Pfam; PF01017; STAT_alpha; 1.	
DR	Pfam; PF02864; STAT_bind; 1.	
DR	Pfam; PF02865; STAT_int; 1.	
DR	SMART; SM00252; SH2; 1.	
DR	PROSITE; PS50001; SH2; 1.	
DR	SEQUENCE 749 AA; 87314 MW; FF7EA8C1734F7C99 CRC64;	
DR	Query Match 100.0%; Score 46; DB 2; Length 749;	
DR	Best Local Similarity 100.0%; Pred. No. 2.5;	
DR	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps	
DR		
DR	Qy 1 KLQELNYNL 9	
DR		
DR	350 KLQELNYNL 358	
DR		
DR	RESULT 4	
DR	Q8C497	
DR	ID Q8C497 PRELIMINARY; PRT; 749 AA.	
DR	AC Q8C497;	
DR	DT 01-MAR-2003 (TrEMBLrel. 23, Created)	
DR	DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)	
DR	DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DR	DE Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length enriched	
DR	DE library, clone:C230091J07 product:signal transducer and activator of	
DR	DE transcription 1, full insert sequence.	
DR	OS Name=Stat1;	
DR	GN Mus musculus (Mouse).	
DR	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
DR	OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
DR	OX NCBI_TaxID=10090;	
DR	[1]	
DR	SEQUENCE FROM N.A.	
DR	STRAIN=C57BL/6J; TISSUE=Cerebellum;	
DR	RC MEDLINE=57879253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;	
DR	CA Carninci P., Hayashizaki Y.;	
DR	RA "High-efficiency full-length cDNA cloning.";	
DR	RT Meth. Enzymol. 303:19-44(1999).	
DR	[2]	
DR	SEQUENCE FROM N.A.	
DR	STRAIN=C57BL/6J; TISSUE=Cerebellum;	
DR	RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;	
DR	RA RIKEN PANTOM Consortium;	
DR	RT "Functional annotation of a full-length mouse cDNA collection.";	
DR	RT Nature 409:685-690(2001).	
DR	[3]	
DR	SEQUENCE FROM N.A.	
DR	STRAIN=C57BL/6J; TISSUE=Cerebellum;	
DR	RC The PANTOM Consortium;	
DR	RA the RIKEN Genome Exploration Research Group Phase I & II Team;	
DR	RT "Analysis of the mouse transcriptome based on functional annotation of	
DR	RT 60,770 full-length cDNAs.";	
DR	RT Nature 420:563-573(2002).	
DR	[4]	
DR	SEQUENCE FROM N.A.	
DR	STRAIN=C57BL/6J; TISSUE=Cerebellum;	
DR	RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;	
DR	CA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,	
DR	RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;	
DR	RT "Normalization and subtraction of cap-trapper-selected cDNAs to	
DR	RT prepare full-length cDNA libraries for rapid discovery of new genes.";	
DR	RT Genome Res. 10:1617-1630(2000).	
DR	[5]	
DR	SEQUENCE FROM N.A.	



```

Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9
Db 350 KQELNYNL 358

RESULT 6
Q9QXK0
ID Q9QXK0 PRELIMINARY; PRT; 749 AA.
AC Q9QXK0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Signal transducer and activator of transcription 1.
GN Name=Stat1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20576429; PubMed=11024034; DOI=10.1074/jbc.M008330200;
RA Chen G., Hohmeier H.E., Newgard C.B.;
RT "Expression of the transcription factor STAT-1 alpha in insulinoma
RT cells protects against cytotoxic effects of multiple cytokines.";
RL J. Biol. Chem. 276:766-772(2001).
DR EMBL; AF205604; AAF20200.1; -.
DR HSSP; P42224; 1BF5.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53 like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF000017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2_1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 749 AA; 87234 MW; D37C634215DED355 CRC64;

Query Match 100.0%; Score 46; DB 2; Length 749;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9
Db 350 KQELNYNL 358

RESULT 7
STAL_HUMAN
ID STAL_HUMAN STANDARD; PRT; 750 AA.
AC P42224;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and activator of transcription 1-alpha/beta
DE (Transcription factor ISGF-3 components p91/p84).
GN Name=STAT1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.; AND SEQUENCE OF 514-524; 654-660 AND 667-672.
RX MEDLINE=92366557; PubMed=1502203;
RA Schindler C., Fu X.-Y., Improtta T., Aebersold R., Darnell J.E. Jr.;

```

```

RT "Proteins of transcription factor ISGF-3: one gene encodes the 91-and
RT 84-kDa ISGF-3 proteins that are activated by interferon alpha.";
RT Proc. Natl. Acad. Sci. U.S.A. 89:7836-7839(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM BETA).
RX TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=95192056; PubMed=7885841;
RA Yan R., Qureshi S., Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "The genomic structure of the STAT genes: multiple exons in coincident
RT sites in Stat1 and Stat2.";
RL Nucleic Acids Res. 23:459-463(1995).
RN [4]
RP PHOSPHORYLATION SITE TYR-701.
RX MEDLINE=95386533; PubMed=7657660; DOI=10.1074/jbc.270.35.20775;
RA Quelle F.W., Thierfelder W., Witthuhn B.A., Tang B., Cohen S.,
RA Ihle J.N.;
RT "Phosphorylation and activation of the DNA binding activity of
RT purified Stat1 by the Janus protein-tyrosine kinases and the epidermal
RT growth factor receptor.";
RL J. Biol. Chem. 270:20775-20780(1995).
RN [5]
RP PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.
RX MEDLINE=95354205; PubMed=7543024; DOI=10.1016/0092-8674(95)90311-9;
RA Wen Z., Zhong Z., Darnell J.E. Jr.;
RT "Maximal activation of transcription by Stat1 and Stat3 requires both
RT tyrosine and serine phosphorylation.";
RL Cell 82:241-250(1995).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 136-710.
RX MEDLINE=98292180; PubMed=9630226; DOI=10.1016/S0092-8674(00)81443-9;
RA Chen X., Vinkemeier U., Zhao Y., Jeruzalmi D., Darnell J.E. Jr.,
RA Xuriyan J.;
RT "Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to
RT DNA.";
RL Cell 93:827-839(1998).
RN [7]
RP VARIANT STAT1 DEFICIENCY SER-706.
RX PubMed=11452125; DOI=10.1126/science.1061154;
RA Dupuis S., Dargemont C., Fieschi C., Thomassin N., Rosenzweig S.,
RA Harris J., Holland S.M., Schreiber R.D., Casanova J.L.;
RT "Impairment of mycobacterial but not viral immunity by a germline
RT human STAT1 mutation.";
RL Science 293:300-303(2001).
RN [8]
RP VARIANT STAT1 DEFICIENCY PRO-600.
RX PubMed=12590259; DOI=10.1038/ng1097;
RA Dupuis S., Jouanguy E., Al-Hajjar S., Fieschi C., Al-Mohsen I.Z.,
RA Al-Jumaa S., Yang K., Chappier A., Eidschinken C., Bid P.,
RA Al-Ghonaim A., Tufenkeji H., Frayha H., Al-Gazlan S., Al-Rayaes H.,
RA Schreiber R.D., Gresser I., Casanova J.L.;

```

RT "Impaired response to interferon-alpha/beta and lethal viral disease  
 RL in human STAT1 deficiency";  
 CC Nat. Genet. 33:388-391(2003).  
 CC -|- FUNCTION: Transcription factor that binds to the IFN-stimulated  
 CC response element (ISRE) and to the GAS element. This multiprotein  
 CC transcription factor is termed ISGF3.  
 CC -|- SUBUNIT: In response to IFN alpha/beta, three subunits (STAT1-  
 CC alpha, STAT1-beta, STAT2) of ISGF3, become phosphorylated on  
 CC tyrosine, migrate into the nucleus, and assemble into a complex  
 CC together with ISGF3 gamma (p48), a DNA-binding protein that  
 CC specifically binds to the IFN-stimulated response element. In  
 CC response to IFN gamma, STAT1 forms homodimers, that also  
 CC translocate into the nucleus to activate IFN gamma-responsive  
 CC genes. Interacts with NMI.  
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus  
 CC in response to phosphorylation.  
 CC -|- ALTERNATIVE PRODUCTS:  
 CC Name=Alpha; Synonyms=p91; Named isoforms=2;  
 CC IsoId=p42224-1; Sequence=Displayed;  
 CC Name=Beta; Synonyms=p84;  
 CC IsoId=p42224-2; Sequence=VSP\_006282;  
 CC -|- PTM: Tyrosine phosphorylated in response to IFN-gamma, IFN-alpha,  
 CC PGIF, and EGF. Serine phosphorylation is also required for maximal  
 CC transcriptional activity (lacking in beta form).  
 CC -|- DISEASE: Defects in STAT1 are the cause of STAT1 deficiency  
 CC [MIM:600555, 209950]. Patients generally suffer from mycobacterial  
 CC or viral diseases. In the case of complete deficiency, patients  
 CC can die of viral disease.  
 CC -|- SIMILARITY: Belongs to the transcription factor STAT family.  
 CC -|- SIMILARITY: Contains 1 SH2 domain.  
 CC -|- DATABASE: NAME=Statbase; NOTE=STAT1 mutation db;  
 CC WWW="http://bioinf.uta.fi/STAT1base/".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M97935; AAB64012.1; -;  
 DR EMBL; M97936; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; BC02704; AA02704.1; -;  
 DR EMBL; U18662; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18663; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18664; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18665; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18666; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18667; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18668; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18669; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18670; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A46159; A46159\_-;  
 DR PDB; 1BF5; X-ray; A=136-710.  
 DR TRANSFAC; T01492; -;  
 DR TRANSFAC; T01573; -;  
 DR Genew; HGNC:11362; STAT1.  
 DR H-InvDB; HIX0002682; -;  
 DR MIM; 600555; -;  
 DR MIM; 209950; -;  
 DR GO; GO:0005737; C:cytoplasm; TAS.  
 DR GO; GO:0005634; C:nucleus; TAS.  
 DR GO; GO:0005062; F:hematopoietin/interferon-class (D200-domain. . .); TAS.  
 DR GO; GO:0003700; F:transcription factor activity; TAS.  
 DR GO; GO:0006915; P:casease activation; TAS.  
 DR GO; GO:0007249; P:I-kappaB kinase/NF-kappaB cascade; TAS.  
 DR GO; GO:0000074; P:regulation of cell cycle; TAS.  
 DR GO; GO:0009613; P:response to pest/pathogen/parasite; TAS.  
 DR GO; GO:0007165; P:signal transduction; TAS.  
 DR GO; GO:0007261; P:STAT protein dimerization; TAS.  
 DR GO; GO:0006366; P:transcription from Pol II promoter; TAS.

DR GO; GO:0007260; P:tyrosine phosphorylation of STAT protein; TAS.  
 DR InterPro; IPR008967; P53\_like\_DNA\_bnd.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001217; STAT.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF01017; STAT\_alpha; 1.  
 DR Pfam; PF02864; STAT\_bind; 1.  
 DR Pfam; PF02865; STAT\_int; 1.  
 DR PROSITE; PS00001; SH2; 1.  
 KW 3D-structure; Activator; Alternative splicing;  
 KW Direct protein sequencing; Disease mutation; DNA-binding;  
 KW Nuclear protein; Phosphorylation; SH2 domain;  
 KW Transcription regulation.  
 FT DOMAIN 573 670 SH2  
 FT MOD\_RES 701 701 Phosphotyrosine (by JAK).  
 FT MOD\_RES 727 727 Missing (in isoform Beta).  
 FT VARSPLIC 713 750 /FTId=VSP\_006282.  
 FT VARIANT 600 600 L->P (in STAT1 deficiency; complete).  
 FT VARIANT 706 706 L->S (in STAT1 deficiency; partial.  
 FT Susceptibility to mycobacterial but not  
 FT to viral disease. Loss of GAP and ISGF3  
 FT activation. Impairs the nuclear  
 FT accumulation of GAP but not of ISGF3 in  
 FT heterozygous cells stimulated by IFNs).  
 FT /FTId=VAR\_018266.  
 FT S->A; Decreased transcriptional  
 FT activation.  
 FT MUTAGEN 727 727  
 FT HELIX 137 179  
 FT TURN 180 181  
 FT HELIX 198 247  
 FT TURN 248 249  
 FT TURN 256 256  
 FT HELIX 257 286  
 FT TURN 290 291  
 FT HELIX 293 316  
 FT STRAND 317 321  
 Query Match 100.0%; Score 46; DB 1; Length 750;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KIQELNYNL 9  
 Db 350 KIQELNYNL 358  
 RESULT 8  
 Q68D00 PRELIMINARY; PRT; 750 AA.  
 ID Q68D00  
 AC Q68D00;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein DKFZp686B04100.  
 GN Name=DKFZp686B04100;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Retina;  
 RG The German cDNA Consortium;  
 RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,  
 RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;  
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBAJ databases.  
 DR EMBL; CR749636; CAHI8430.1; -;  
 DR InterPro; IPR008967; P53\_like\_DNA\_bnd.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001217; STAT.  
 DR Pfam; PF00017; SH2; 1.



```

DR Pfam; PF01017; STAT_alpha; 2.
DR Pfam; PF02864; STAT_bind; 2.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SMO0252; SH2; 2.
DR PROSITE; PS00001; SH2; 2.
SQ SEQUENCE 1165 AA; 134606 MW; 75C8B983B5D3DE7 CRC64;

Query Match 100.0%; Score 46; DB 2; Length 1165;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KLOELNYNL 9
Db 781 KLOELNYNL 789

RESULT 11
O764M5 PRELIMINARY; PRT; 757 AA.
AC Q764M5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Signal transducer and activator of transcription 1.
GN Name=stat1;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Uenishi H., Eguchi T., Suzuki K., Sawasaki T., Toki D., Shinkai H.,
RA Okumura N., Hamashima N., Awata T.;
RT "PDEE (Pig EST Data Explorer): construction of a database for ESTs
RT derived from porcine full-length cDNA libraries.";
RL Nucleic Acids Res. 32:D484-D488(2004).
DR EMBL; AB116564; BAD06318.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SMO0252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
SQ SEQUENCE 757 AA; 88166 MW; B6093218F2F6A029 CRC64;

Query Match 91.3%; Score 42; DB 2; Length 757;
Best Local Similarity 88.9%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KLOELNYNL 9
Db 350 KLOELNYNL 358

RESULT 12
Q8JGNO PRELIMINARY; PRT; 751 AA.
AC Q8JGNO;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Stc1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.

```

```

OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21098508; PubMed=11164887; DOI=10.1016/S0145-305X(00)00050-1;
RA Turpen J.B., Carlson D.L., Huang C.;
RA "Cloning and developmental expression of Xenopus Stat1.";
RA Dev. Comp. Immunol. 25:219-229(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Turpen J.B., Carlson D.L., Huang C.;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY101602; AAM51552.1; -.
DR HSSP; P42224; 1BF5.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SMO0252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
SQ SEQUENCE 751 AA; 86497 MW; 788810A08B0889EA CRC64;

Query Match 87.0%; Score 40; DB 2; Length 751;
Best Local Similarity 88.9%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KLOELNYNL 9
Db 350 KLOELNYNL 358

RESULT 13
Q97169 PRELIMINARY; PRT; 354 AA.
AC Q97169;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE DNA uptake protein.
GN OrderedLocusNames=CAC1784;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RX DOI=10.1128/JB.183.16.4823-4838.2001;
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Nosling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007687; AAK79749.1; -.
DR FIR; B97120; B97120.
DR GO; GO:0009294; P:DNA mediated transformation; IEA.
DR InterPro; IPR003488; SMF.
DR Pfam; PF02481; SMF; 1.
DR TIGSFAMS; TIGR00732; dprA; 1.
KW Complete proteome.
SQ SEQUENCE 354 AA; 40570 MW; 105FDB089BFCD8C2 CRC64;

Query Match 82.6%; Score 38; DB 2; Length 354;

```

Best Local Similarity 77.8%; Pred. No. 44;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLOELNYNL 9  
95 KLNELNYNV 103

## RESULT 14

QJUMC3  
ID Q8JMC3 PRELIMINARY; PRT; 115 AA.  
AC Q8JMC3;  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
DE Hypothetical protein.  
OS Mamestra configurata nucleopolyhedrovirus B.  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=204440;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22080411; PubMed=12083822; DOI=10.1006/viro.2002.1411;  
RA Li L., Donly C., Li Q., Willis L.G., Keddie B.A., Erlandson M.A.,  
RA Theilmann D.A.;  
RT "Identification and genomic analysis of a second species of  
RT nucleopolyhedrovirus isolated from Mamestra configurata.";  
RL Virology 297:226-244 (2002).  
DR EMBL; AY126275; AAM95016.1; --  
KW Hypothetical protein.  
SQ SEQUENCE 115 AA; 13082 MW; 0D35176637F6A7D2 CRC64;

Query Match 78.3%; Score 36; DB 2; Length 115;  
Best Local Similarity 75.0%; Pred. No. 33;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LOELNYNL 9  
103 IOELNYNV 110

## RESULT 15

Q8QLK0  
ID Q8QLK0 PRELIMINARY; PRT; 117 AA.  
AC Q8QLK0;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mamestra configurata nucleopolyhedrovirus (MaconPV).  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=191492;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=90/2;  
RX MEDLINE=97163493; PubMed=9010313;  
RA Li S., Erlandson M., Moody D., Gillott C.;  
RT "A physical map of the Mamestra configurata nucleopolyhedrovirus  
RT genome and sequence analysis of the polyhedrin gene.";  
RL J. Gen. Virol. 78:265-271 (1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=90/2;  
RX MEDLINE=21884635; PubMed=1186270; DOI=10.1006/viro.2001.1313;  
RA Li Q., Donly C., Li L., Willis L.G., Theilmann D.A., Erlandson M.;  
RT "Sequence and organization of the Mamestra configurata  
RT nucleopolyhedrovirus genome.";  
RL Virology 294:106-121 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=90/2;  
RA Li Q., Donly C., Li L., Willis L.G., Theilmann D.A., Erlandson M.A.;

RL Submitted (JAN-2002) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; U59461; AAM09142.1; --  
KW Hypothetical protein.  
SQ SEQUENCE 117 AA; 13296 MW; E9E0E3DAD6ED724A CRC64;

Query Match 78.3%; Score 36; DB 2; Length 117;  
Best Local Similarity 75.0%; Pred. No. 33;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LOELNYNL 9  
105 IOELNYNV 112

Search completed: November 18, 2005, 01:08:26  
Job time : 44.9677 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 00:48:52 ; Search time 13.6452 Seconds  
(without alignments)  
49.237 Million cell updates/sec

Title: US-10-006-177-10

Perfect score: 46

Sequence: 1 KQELNYNL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pap:\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pap:\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pap:\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pap:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pap:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	171	3	US-09-087-465-18
2	46	100.0	268	3	US-09-387-418A-12
3	46	100.0	582	4	US-09-430-806A-3
4	46	100.0	712	1	US-08-369-796-6
5	46	100.0	712	2	US-08-852-091-6
6	46	100.0	712	2	US-08-820-754-6
7	46	100.0	712	3	US-08-956-652-6
8	46	100.0	712	3	US-08-956-869-6
9	46	100.0	712	3	US-08-948-547-6
10	46	100.0	712	3	US-08-956-653A-6
11	46	100.0	712	4	US-08-212-185-6
12	46	100.0	712	4	US-09-430-806A-2
13	46	100.0	712	5	PCT-US95-17025-6
14	46	100.0	729	4	US-09-917-254-97
15	46	100.0	740	1	US-08-276-039A-12
16	46	100.0	740	1	US-08-781-890-12
17	46	100.0	750	1	US-08-369-796-4
18	46	100.0	750	2	US-08-852-091-4
19	46	100.0	750	2	US-08-820-754-4
20	46	100.0	750	3	US-08-956-652-4
21	46	100.0	750	3	US-08-956-869-4
22	46	100.0	750	3	US-08-948-547-4
23	46	100.0	750	3	US-09-087-465-2
24	46	100.0	750	3	US-09-364-970-1
25	46	100.0	750	3	US-09-364-970-8
26	46	100.0	750	3	US-08-956-653A-4
27	46	100.0	750	4	US-09-972-800A-2

28	46	100.0	750	4	US-08-212-185-4	Sequence 4, Appli
29	46	100.0	750	4	US-09-430-806A-1	Sequence 1, Appli
30	46	100.0	750	5	PCT-US95-17025-4	Sequence 4, Appli
31	46	100.0	767	4	US-09-949-016-8350	Sequence 8350, Ap
32	36	78.3	487	4	US-09-583-110-4658	Sequence 4658, Ap
33	36	78.3	490	4	US-09-107-433-4640	Sequence 4640, Ap
34	35	76.1	196	3	US-09-129-030-2	Sequence 2, Appli
35	35	76.1	333	4	US-09-583-110-4020	Sequence 4020, Ap
36	35	76.1	345	4	US-09-248-796A-19172	Sequence 19172, A
37	35	76.1	361	4	US-09-107-433-5205	Sequence 5205, Ap
38	34	73.9	212	4	US-09-270-767-48155	Sequence 48155, A
39	34	73.9	218	3	US-09-134-001C-5032	Sequence 5032, Ap
40	34	73.9	306	4	US-09-107-532A-5918	Sequence 5918, Ap
41	34	73.9	348	4	US-09-134-000C-5396	Sequence 5396, Ap
42	34	73.9	660	4	US-09-248-796A-19966	Sequence 19966, A
43	33	71.7	244	4	US-09-328-352-5597	Sequence 5597, A
44	33	71.7	250	4	US-09-248-796A-27542	Sequence 27542, A
45	33	71.7	566	4	US-09-411-628-4	Sequence 4, Appli

## ALIGNMENTS

### RESULT 1

US-09-087-465-18  
; Sequence 18, Application US/09087465A  
; Patent No. 6160092  
; GENERAL INFORMATION:  
; APPLICANT: Vinkemeier, Uwe  
; APPLICANT: Chen, Xiaomin  
; APPLICANT: Darnell Jr., James E  
; APPLICANT: Kuriyan, John  
; TITLE OF INVENTION: A CRYSTAL OF THE CORE PORTION OF A STAT AND METHODS OF  
; FILE REFERENCE: 600-1-229  
; CURRENT APPLICATION NUMBER: US/09/087,465A  
; CURRENT FILING DATE: 1998-05-29  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-087-465-18

Query Match 100.0%; Score 46; DB 3; Length 171;  
Best Local Similarity 100.0%; Pred. No. 0.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9  
Db 34 KQELNYNL 42

### RESULT 2

US-09-387-418A-12  
; Sequence 12, Application US/09387418A  
; Patent No. 6391572  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Xiaokui  
; APPLICANT: Wrzeszczynska, Melissa H  
; APPLICANT: Horvath, Curt M  
; APPLICANT: Darnell Jr., James E  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR  
; FILE REFERENCE: 600-1-253  
; CURRENT APPLICATION NUMBER: US/09/387,418A  
; CURRENT FILING DATE: 1999-08-31  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 268  
; TYPE: PRT

```
; ORGANISM: Mus musculus
US-09-387-418A-12

Query Match      100.0%; Score 46; DB 3; Length 268;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      244 KLOELNYNL 252

RESULT 3
US-09-430-806A-3
; Sequence 3, Application US/09430806A
; Patent No. 6720154
; GENERAL INFORMATION:
; APPLICANT: Vinkemeier, Uwe
; APPLICANT: Darnell Jr., James E.
; TITLE OF INVENTION: PURIFIED STAT PROTEINS AND METHODS OF PURIFYING THEREOF
; FILE REFERENCE: 600-1-182 N
; CURRENT APPLICATION NUMBER: US/09/430.806A
; CURRENT FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 08/951,130
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: 60/028,176
; PRIOR FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-430-806A-3

Query Match      100.0%; Score 46; DB 4; Length 582;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      219 KLOELNYNL 227

RESULT 4
US-08-369-796-6
; Sequence 6, Application US/08369796
; Patent No. 5716622
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/369,796
; FILING DATE: 06-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-852-091-6

Query Match      100.0%; Score 46; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 358

RESULT 5
US-08-852-091-6
; Sequence 6, Application US/08852091
; Patent No. 5883228
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,091
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/369,796
; FILING DATE: 06-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-852-091-6

Query Match      100.0%; Score 46; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 358

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 712 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-369-796-6

Query Match      100.0%; Score 46; DB 1; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 358

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 712 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-369-796-6
```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
| | | | |  
Db 350 KLOELNYNL 358

RESULT 6  
US-08-820-754-6  
; Sequence 6, Application US/08820754  
; Patent No. 5976835  
; GENERAL INFORMATION:  
; APPLICANT: Darnell Jr., James E.  
; APPLICANT: Schindler, Christian W.  
; APPLICANT: Fu, Xian-Yuan  
; APPLICANT: Wen, Zilong  
; APPLICANT: Zhong, Zhong  
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/820,754  
; FILING DATE: 19-MAR-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,185  
; FILING DATE: 11-MAR-1994  
; APPLICATION NUMBER: US 07/980,498  
; FILING DATE: 23-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/854,296  
; FILING DATE: 19-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO US93/02569  
; FILING DATE: 19-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/126,588  
; FILING DATE: 24-SEP-1993  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-820-754-6

Query Match 100.0%; Score 46; DB 2; Length 712;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
| | | | |  
Db 350 KLOELNYNL 358

RESULT 7  
US-08-956-652-6  
; Sequence 6, Application US/08956652  
; Patent No. 6013475  
; GENERAL INFORMATION:  
; APPLICANT: Darnell Jr., James E.  
; APPLICANT: Schindler, Christian W.  
; APPLICANT: Fu, Xian-Yuan  
; APPLICANT: Wen, Zilong  
; APPLICANT: Zhong, Zhong  
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/956,652  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,185  
; FILING DATE: 11-MAR-1994  
; APPLICATION NUMBER: US 07/980,498  
; FILING DATE: 23-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/854,296  
; FILING DATE: 19-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO US93/02569  
; FILING DATE: 19-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/126,588  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-956-652-6

Query Match 100.0%; Score 46; DB 3; Length 712;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
| | | | |  
Db 350 KLOELNYNL 358

RESULT 8  
US-08-956-869-6  
; Sequence 6, Application US/08956869

```

; Patent No. 6030808
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,869
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/212,185
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,296
; FILING DATE: 19-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US93/02569
; FILING DATE: 19-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,588
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-956-869-6

```

```

Query Match 100.0%; Score 46; DB 3; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 KLOELNYNL 9
Db 350 KLOELNYNL 358

```

```

RESULT 9
US-08-948-547-6
; Sequence 6, Application US/08948547
; Patent No. 612418
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong

```

```

; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,547
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,185
; FILING DATE: 11-MAR-1994
; APPLICATION NUMBER: US 07/980,498
; FILING DATE: 23-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,296
; FILING DATE: 19-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US93/02569
; FILING DATE: 19-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,588
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-948-547-6

```

```

Query Match 100.0%; Score 46; DB 3; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 KLOELNYNL 9
Db 350 KLOELNYNL 358

```

```

RESULT 10
US-08-956-653A-6
; Sequence 6, Application US/08956653A
; Patent No. 6338949
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson

```

STREET: 411 Hackensack Avenue  
 CITY: Hackensack  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07601  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/956,653A  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/212,185  
 FILING DATE: 11-MAR-1994  
 APPLICATION NUMBER: US 07/980,498  
 FILING DATE: 23-NOV-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/854,296  
 FILING DATE: 19-MAR-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO US93/02569  
 FILING DATE: 19-MAR-1993  
 APPLICATION NUMBER: US 08/126,588  
 FILING DATE: 24-SEP-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jackson Esq., David A.  
 REGISTRATION NUMBER: 26,742  
 REFERENCE/DOCKET NUMBER: 600-1-195  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201 487-5800  
 TELEFAX: 201 343-1684  
 TELEX: 133521  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 712 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-956-653A-6

Query Match 100.0%; Score 46; DB 3; Length 712;  
 Best Local Similarity 100.0%; Pred. No. 0.89;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KJQELNYNL 9  
 Db 350 KJQELNYNL 358

RESULT 11  
 US-08-212-185-6  
 ; Sequence 6, Application US/08212185  
 ; Patent No. 6805442  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Darnell Jr., James E.  
 ; APPLICANT: Schindler, Christian W.  
 ; APPLICANT: Fu, Xian-Yuan  
 ; APPLICANT: Wen, Zilong  
 ; APPLICANT: Zhong, Zhong  
 ; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
 ; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF  
 ; NUMBER OF SEQUENCES: 25  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Klauber & Jackson  
 ; STREET: 411 Hackensack Avenue  
 ; CITY: Hackensack  
 ; STATE: New Jersey  
 ; COUNTRY: USA  
 ; ZIP: 07601

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/212,185  
 ; FILING DATE: 11-MAR-1994  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/980,498  
 ; FILING DATE: 23-NOV-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/854,296  
 ; FILING DATE: 19-MAR-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: WO US93/02569  
 ; FILING DATE: 19-MAR-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/126,588  
 ; FILING DATE: 24-SEP-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Jackson Esq., David A.  
 ; REGISTRATION NUMBER: 26,742  
 ; REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 201 487-5800  
 ; TELEFAX: 201 343-1684  
 ; TELEX: 133521  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 712 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-212-185-6

Query Match 100.0%; Score 46; DB 4; Length 712;  
 Best Local Similarity 100.0%; Pred. No. 0.89;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KJQELNYNL 9  
 Db 350 KJQELNYNL 358

RESULT 12  
 US-09-430-806A-2  
 ; Sequence 2, Application US/09430806A  
 ; Patent No. 6720154  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Vinkemeier, Uwe  
 ; APPLICANT: Darnell Jr., James E.  
 ; TITLE OF INVENTION: PURIFIED STAT PROTEINS AND METHODS OF PURIFYING THEREOF  
 ; FILE REFERENCE: 600-1-182 N  
 ; CURRENT APPLICATION NUMBER: US/09/430,806A  
 ; CURRENT FILING DATE: 1999-11-02  
 ; PRIOR APPLICATION NUMBER: 08/951,130  
 ; PRIOR FILING DATE: 1997-10-15  
 ; PRIOR APPLICATION NUMBER: 60/028,176  
 ; PRIOR FILING DATE: 1996-10-15  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 712  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-430-806A-2

Query Match 100.0%; Score 46; DB 4; Length 712;  
 Best Local Similarity 100.0%; Pred. No. 0.89;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 KLOELNYNL 9
      |||||||
Db     350 KLOELNYNL 358

RESULT 13
PCT-US95-17025-6
; Sequence 6, Application PC/TUS9517025
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Hohvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17025
; FILING DATE: 28-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/369,796
; FILING DATE: 06-JAN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-17025-6

      Query Match      100.0%; Score 46; DB 5; Length 712;
      Best Local Similarity 100.0%; Pred. No. 0.89;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
      |||||||
Db     350 KLOELNYNL 358

RESULT 14
US-09-917-254-97
; Sequence 97, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224 (JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
```

```
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 97
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-917-254-97

      Query Match      100.0%; Score 46; DB 4; Length 729;
      Best Local Similarity 100.0%; Pred. No. 0.91;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
      |||||||
Db     367 KLOELNYNL 375

RESULT 15
US-08-276-099A-12
; Sequence 12, Application US/08276099A
; Patent No. 5591825
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L
; APPLICANT: Hou, Jinzhao
; TITLE OF INVENTION: INTERLEUKIN-4 SIGNAL TRANSDUCERS AND
; TITLE OF INVENTION: BINDING ASSAYS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,099A
; FILING DATE: 15-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59451-1/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 740 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-276-099A-12

      Query Match      100.0%; Score 46; DB 1; Length 740;
      Best Local Similarity 100.0%; Pred. No. 0.93;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
      |||||||
Db     339 KLOELNYNL 347

Search completed: November 18, 2005, 09:36:04
Job time : 15.6452 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 01:08:37 ; Search time 44.3226 Seconds  
(without alignments)  
84.961 Million cell updates/sec

Title: US-10-006-177-10

Perfect score: 46

Sequence: 1 KLOELNYNL 9

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubaa/US09A\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubaa/US09B\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubaa/US09C\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubaa/US09D\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubaa/US09E\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubaa/US10F\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	9	14	US-10-006-177-10
2	46	100.0	9	14	US-10-006-177-11
3	46	100.0	9	16	US-10-473-127-138
4	46	100.0	171	16	US-10-473-127-1555
5	46	100.0	268	13	US-10-090-185-12
6	46	100.0	582	14	US-10-245-120-3
7	46	100.0	582	16	US-10-473-127-1414
8	46	100.0	701	16	US-10-473-127-1407
9	46	100.0	712	11	US-09-876-773-6
10	46	100.0	712	14	US-10-245-120-2
11	46	100.0	712	16	US-10-473-127-1401
					Sequence 10, Appl
					Sequence 11, Appl
					Sequence 138, Appl
					Sequence 1555, Ap
					Sequence 12, Appl
					Sequence 3, Appli
					Sequence 1414, Ap
					Sequence 1407, Ap
					Sequence 6, Appli
					Sequence 1401, Ap.

Query Match 100.0%; Score 46; DB 14; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.7e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9

Db 1 KLOELNYNL 9

#### ALIGNMENTS

##### RESULT 1

US-10-006-177-10  
; Sequence 10, Application US/10006177  
; Publication No. US20030165513A1  
; GENERAL INFORMATION:  
; APPLICANT: Ramakrishna, Venky  
; APPLICANT: Ross, Mark  
; APPLICANT: Philip, Ramila  
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treat  
; FILE REFERENCE: 26747-35  
; CURRENT APPLICATION NUMBER: US/10/006.177  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: US/60/251,022  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: US/60/256,824  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Epitopic Peptide  
US-10-006-177-10

12	46	100.0	712	16	US-10-473-127-1409	Sequence 1409, Ap
13	46	100.0	712	16	US-10-473-127-1411	Sequence 1411, Ap
14	46	100.0	712	16	US-10-473-127-1413	Sequence 1413, Ap
15	46	100.0	712	17	US-10-639-617-6	Sequence 6, Appli
16	46	100.0	712	17	US-10-936-390-5	Sequence 5, Appli
17	46	100.0	739	16	US-10-473-127-1406	Sequence 1406, Ap
18	46	100.0	749	9	US-09-833-205-4	Sequence 4, Appli
19	46	100.0	750	9	US-09-833-205-2	Sequence 2, Appli
20	46	100.0	750	11	US-09-876-773-4	Sequence 4, Appli
21	46	100.0	750	14	US-10-245-120-1	Sequence 1, Appli
22	46	100.0	750	14	US-10-308-279-44	Sequence 44, Appl
23	46	100.0	750	16	US-10-755-889-352	Sequence 352, App
24	46	100.0	750	16	US-10-755-889-823	Sequence 823, App
25	46	100.0	750	16	US-10-473-127-1402	Sequence 1402, Ap
26	46	100.0	750	16	US-10-473-127-1403	Sequence 1403, Ap
27	46	100.0	750	16	US-10-473-127-1408	Sequence 1408, Ap
28	46	100.0	750	16	US-10-473-127-1410	Sequence 1410, Ap
29	46	100.0	750	16	US-10-473-127-1412	Sequence 1412, Ap
30	46	100.0	750	16	US-10-473-127-1415	Sequence 1415, Ap
31	46	100.0	750	16	US-10-473-127-1416	Sequence 1416, Ap
32	46	100.0	750	17	US-10-492-043-19	Sequence 19, Appli
33	46	100.0	750	17	US-10-639-617-4	Sequence 4, Appli
34	46	100.0	750	18	US-10-491-545A-55	Sequence 55, Appl
35	46	100.0	750	18	US-10-631-467-658	Sequence 658, App
36	46	100.0	786	9	US-09-925-237-550	Sequence 550, App
37	46	100.0	786	16	US-10-473-127-1405	Sequence 1405, Ap
38	38	82.6	459	18	US-10-450-763-32710	Sequence 32710, A
39	36	78.3	197	15	US-10-108-260A-2644	Sequence 2644, Ap
40	36	78.3	441	15	US-10-108-260A-3815	Sequence 3815, Ap
41	36	78.3	487	17	US-10-472-928-880	Sequence 880, App
42	36	78.3	490	18	US-10-617-320-4640	Sequence 4640, Ap
43	35	76.1	240	17	US-10-732-923-16841	Sequence 16841, A
44	35	76.1	300	15	US-10-424-599-226964	Sequence 226964,
45	35	76.1	355	17	US-10-472-928-490	Sequence 490, App

RESULT 2  
US-10-006-177-11  
; Sequence 11, Application US/10006177  
; Publication No. US20030165513A1  
; GENERAL INFORMATION:  
; APPLICANT: Ramakrishna, Venky  
; APPLICANT: Rose, Mark  
; APPLICANT: Philip, Ramila  
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treatment  
; FILE REFERENCE: 26747-35  
; CURRENT APPLICATION NUMBER: US/10/006,177  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: US/60/251,022  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: US/60/256,824  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Epitopic Peptide  
US-10-006-177-11

Query Match 100.0%; Score 46; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLQELNYNL 9  
|||  
Db 1 KLQELNYNL 9

RESULT 3  
US-10-473-127-138  
; Sequence 138, Application US/10473127  
; Publication No. US20040236091A1  
; GENERAL INFORMATION:  
; APPLICANT: Zycos Inc.  
; TITLE OF INVENTION: TRANSLATIONAL PROFILING  
; FILE REFERENCE: 08191-026W01  
; CURRENT APPLICATION NUMBER: US/10/473,127  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 60/279,495  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: 60/292,544  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/310,801  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/326,370  
; PRIOR FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: 60/336,780  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/358,985  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 2041  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 138  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-473-127-138

Query Match 100.0%; Score 46; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLQELNYNL 9  
|||

Db 1 KLQELNYNL 9  
|||  
RESULT 4  
US-10-473-127-1555  
; Sequence 1555, Application US/10473127  
; Publication No. US20040236091A1  
; GENERAL INFORMATION:  
; APPLICANT: Zycos Inc.  
; TITLE OF INVENTION: TRANSLATIONAL PROFILING  
; FILE REFERENCE: 08191-026W01  
; CURRENT APPLICATION NUMBER: US/10/473,127  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 60/279,495  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: 60/292,544  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/310,801  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/326,370  
; PRIOR FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: 60/336,780  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/358,985  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 2041  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1555  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-473-127-1555

Query Match 100.0%; Score 46; DB 16; Length 171;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLQELNYNL 9  
|||  
Db 34 KLQELNYNL 42

RESULT 5  
US-10-090-185-12  
; Sequence 12, Application US/10090185  
; Publication No. US20020197647A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Xiaokui  
; APPLICANT: Wzrzeszynska, Melissa H  
; APPLICANT: Horvath, Curt M  
; APPLICANT: Darnell Jr., James E  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P  
; FILE REFERENCE: 600-1-253  
; CURRENT APPLICATION NUMBER: US/10/090,185  
; CURRENT FILING DATE: 2002-03-04  
; PRIOR FILING DATE: 1999-08-31  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 268  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-090-185-12

Query Match 100.0%; Score 46; DB 13; Length 268;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLQELNYNL 9  
|||



```
Db      244 KLQELNYNL 252

RESULT 6
US-10-245-120-3
; Sequence 3, Application US/10245120
; Publication No. US20030092066A1
; GENERAL INFORMATION:
; APPLICANT: Vinkemeier, Uwe
; TITLE OF INVENTION: PURIFIED STAT PROTEINS AND METHODS OF PURIFYING THEREOF
; FILE REFERENCE: 600-1-182 N
; CURRENT APPLICATION NUMBER: US/10/245,120
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US/09/430,806
; PRIOR FILING DATE: 1998-11-02
; PRIOR APPLICATION NUMBER: 08/951,130
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: 60/028,176
; PRIOR FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-245-120-3

Query Match      100.0%; Score 46; DB 14; Length 582;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KLQELNYNL 9
Db      219 KLQELNYNL 227

RESULT 7
US-10-473-127-1414
; Sequence 1414, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1414
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1414

Query Match      100.0%; Score 46; DB 16; Length 582;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KLQELNYNL 9
Db      219 KLQELNYNL 227

RESULT 8
US-10-473-127-1407
; Sequence 1407, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1407
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1407

Query Match      100.0%; Score 46; DB 16; Length 701;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KLQELNYNL 9
Db      339 KLQELNYNL 347

RESULT 9
US-09-876-773-6
; Sequence 6, Application US/09876773
; Publication No. US20040058318A1
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; Schindler, Christian W.
; Fu, Xian-Yuan
; Wen, Zilong
; Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/876,773
; FILING DATE: 07-Jun-2001
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,185
```

; FILING DATE: 11-MAR-1994  
; APPLICATION NUMBER: US 07/980,498  
; FILING DATE: 23-NOV-1992  
; APPLICATION NUMBER: US 07/854,296  
; FILING DATE: 19-MAR-1992  
; APPLICATION NUMBER: WO US93/02569  
; FILING DATE: 19-MAR-1993  
; APPLICATION NUMBER: US 08/126,588  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-876-773-6

Query Match 100.0%; Score 46; DB 11; Length 712;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIQELNYNL 9  
|||  
Db 350 KIQELNYNL 358

RESULT 10  
US-10-245-120-2  
; Sequence 2, Application US/10245120  
; Publication No. US20030092066A1  
; GENERAL INFORMATION:  
; APPLICANT: Vinkemeier, Uwe  
; TITLE OF INVENTION: PURIFIED STAT PROTEINS AND METHODS OF PURIFYING THEREOF  
; FILE REFERENCE: 600-1-182 N  
; CURRENT APPLICATION NUMBER: US/10/245,120  
; CURRENT FILING DATE: 2002-09-17  
; PRIOR APPLICATION NUMBER: US/09/430,806  
; PRIOR FILING DATE: 1999-11-02  
; PRIOR APPLICATION NUMBER: 08/951,130  
; PRIOR FILING DATE: 1997-10-15  
; PRIOR APPLICATION NUMBER: 60/028,176  
; PRIOR FILING DATE: 1996-10-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 712  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-245-120-2

Query Match 100.0%; Score 46; DB 14; Length 712;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIQELNYNL 9  
|||  
Db 350 KIQELNYNL 358

RESULT 11  
US-10-473-127-1401  
; Sequence 1401, Application US/10473127  
; Publication No. US20040236091A1

; GENERAL INFORMATION:  
; APPLICANT: Zycos Inc.  
; TITLE OF INVENTION: TRANSLATIONAL PROFILING  
; FILE REFERENCE: 08191-026W01  
; CURRENT APPLICATION NUMBER: US/10/473,127  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 60/279,495  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: 60/292,544  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/310,801  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/326,370  
; PRIOR FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: 60/336,780  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/358,985  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 2041  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1401  
; LENGTH: 712  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-473-127-1401

Query Match 100.0%; Score 46; DB 16; Length 712;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIQELNYNL 9  
|||  
Db 350 KIQELNYNL 358

RESULT 12  
US-10-473-127-1409  
; Sequence 1409, Application US/10473127  
; Publication No. US20040236091A1  
; GENERAL INFORMATION:  
; APPLICANT: Zycos Inc.  
; TITLE OF INVENTION: TRANSLATIONAL PROFILING  
; FILE REFERENCE: 08191-026W01  
; CURRENT APPLICATION NUMBER: US/10/473,127  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 60/279,495  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: 60/292,544  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/310,801  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/326,370  
; PRIOR FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: 60/336,780  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/358,985  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 2041  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1409  
; LENGTH: 712  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-473-127-1409

Query Match 100.0%; Score 46; DB 16; Length 712;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIQELNYNL 9  
|||  
Db 350 KIQELNYNL 358

## RESULT 13

US-10-473-127-1411  
; Sequence 1411, Application US/10473127  
; Publication No. US20040236091A1  
; GENERAL INFORMATION:  
; APPLICANT: Zycos Inc.  
; TITLE OF INVENTION: TRANSLATIONAL PROFILING  
; FILE REFERENCE: 08191-026W01  
; CURRENT APPLICATION NUMBER: US/10/473,127  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 60/279,495  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: 60/292,544  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/310,801  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/326,370  
; PRIOR FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: 60/336,780  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/358,985  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 2041  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1411  
; LENGTH: 712  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-473-127-1411

Query Match 100.0%; Score 46; DB 16; Length 712;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## Qy

1 KLOELNYNL 9

|||||

350 KLOELNYNL 358

## RESULT 14

US-10-473-127-1413  
; Sequence 1413, Application US/10473127  
; Publication No. US20040236091A1  
; GENERAL INFORMATION:  
; APPLICANT: Zycos Inc.  
; TITLE OF INVENTION: TRANSLATIONAL PROFILING  
; FILE REFERENCE: 08191-026W01  
; CURRENT APPLICATION NUMBER: US/10/473,127  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 60/279,495  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: 60/292,544  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/310,801  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/326,370  
; PRIOR FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: 60/336,780  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/358,985  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 2041  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1413  
; LENGTH: 712  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-473-127-1413

Query Match 100.0%; Score 46; DB 16; Length 712;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## Qy

1 KLOELNYNL 9

|||||

350 KLOELNYNL 358

## RESULT 15

US-10-639-617-6  
; Sequence 6, Application US/10639617  
; Publication No. US20050079543A1  
; GENERAL INFORMATION:  
; APPLICANT: Darnell Jr., James E.  
; Schindler, Christian W.  
; Fu, Xian-Yuan  
; Wen, Zilong  
; Zhong, Zhong  
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
SEQUENCES AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/639,617  
; FILING DATE: 12-Aug-2003  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/212,185  
; FILING DATE: 11-MAR-1994  
; APPLICATION NUMBER: US 07/980,498  
; FILING DATE: 23-NOV-1992  
; APPLICATION NUMBER: US 07/854,296  
; FILING DATE: 19-MAR-1992  
; APPLICATION NUMBER: WO US93/02569  
; FILING DATE: 19-MAR-1993  
; APPLICATION NUMBER: US 08/126,588  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-639-617-6

Query Match 100.0%; Score 46; DB 17; Length 712;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## Qy

1 KLOELNYNL 9

|||||

350 KLOELNYNL 358

Search completed: November 18, 2005, 10:44:45

Job time : 45.3226 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: November 17, 2005, 22:03:06 ; Search time 48.7742 Seconds  
(without alignments)  
71.366 Million cell updates/sec

Title: US-10-006-177-11

Perfect score: 46

Sequence: 1 KLELNLYNL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*

1: geneseqp1980s.\*

2: geneseqp1990s.\*

3: geneseqp2000s.\*

4: geneseqp2001s.\*

5: geneseqp2002s.\*

6: geneseqp2003as.\*

7: geneseqp2003bs.\*

8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	9	5	ABG32311 Immunogen
2	46	100.0	9	5	ABG32310 HLA-A2 as
3	46	100.0	9	6	ABU03358 Human exp
4	46	100.0	171	4	AAB19969 Human STA
5	46	100.0	171	6	ABU04889 Human exp
6	46	100.0	268	4	AAV72844 Mouse Sta
7	46	100.0	582	2	AAW62996 Human tru
8	46	100.0	582	6	ABU04748 Human exp
9	46	100.0	582	8	ADH57036 Truncated
10	46	100.0	680	6	ABR59713 Human sig
11	46	100.0	701	2	AAK41335 84 kD ISG
12	46	100.0	701	6	ABU04741 Human exp
13	46	100.0	712	2	AAW72079 Human Sta
14	46	100.0	712	2	AAW03170 Human STA
15	46	100.0	712	2	AAW62995 Human Sta
16	46	100.0	712	6	ABU04747 Human exp
17	46	100.0	712	6	ABU04735 Human exp
18	46	100.0	712	6	ABU04745 Human exp
19	46	100.0	712	6	ABU04743 Human exp
20	46	100.0	712	8	ADH57035 Truncated
21	46	100.0	712	8	ADJ32352 Human STA
22	46	100.0	712	8	ADN04464 Antipsoi
23	46	100.0	729	2	AAU84356 Protein S
24	46	100.0	739	2	AAU841334 91 kD ISG
25	46	100.0	739	6	ABU04740 Human exp

26	46	100.0	749	5	AAG78526 Rat STAT-
27	46	100.0	750	2	AAR72078 Human Sta
28	46	100.0	750	2	AAW03168 Human STA
29	46	100.0	750	2	AAW62994 Human Sta
30	46	100.0	750	4	AAB19962 Human sig
31	46	100.0	750	5	AAE15172 Human Sta
32	46	100.0	750	5	AAG78525 Human STA
33	46	100.0	750	6	ABP98871 Human STA
34	46	100.0	750	6	ABU04749 Human exp
35	46	100.0	750	6	ABU04742 Human exp
36	46	100.0	750	6	ABU04744 Human exp
37	46	100.0	750	6	ABU04737 Human exp
38	46	100.0	750	6	ABU04736 Human exp
39	46	100.0	750	6	ABU04750 Human exp
40	46	100.0	750	6	ABU04746 Human exp
41	46	100.0	750	7	ADE34560 Human tra
42	46	100.0	750	8	ADH57034 Human Sta
43	46	100.0	750	8	ADJ32351 Human STA
44	46	100.0	750	8	ADJ75406 Marker ge
45	46	100.0	750	8	ADN03678 Antipsoi

## ALIGNMENTS

### RESULT 1

ABG32311  
ID ABG32311 standard; peptide; 9 AA.

AC ABG32311;

DT 05-NOV-2002 (first entry)

XX Immunogenic peptide from human signal transducer and activator protein.

Human; immunogen; epitope; HLA-A1; human leukocyte antigen; CTL;

cytotoxic Y lymphocyte; cytostatic; cancer; colorectal carcinoma;

ovarian carcinoma; lung carcinoma; prostate carcinoma; vaccine; tumour;

HLA-2; passive immunotherapy; signal transducer and activator protein.

OS Homo sapiens.

XX WO200246416-A2.

XX 13-JUN-2002.

XX 04-DEC-2001; 2001WO-US047290.

XX 04-DEC-2000; 2000US-0251022P.

XX 20-DEC-2000; 2000US-0256824P.

XX (ARGO-) ARGONEX INC.

XX Ramakrishna V, Ross M, Philip R;

XX WPI; 2002-619021/66.

XX New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte, and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.

XX Claim 1; Page 50; 60pp; English.

XX The invention relates to an immunogen comprising an isolated polypeptide whose amino acid sequence comprises an epitopic peptide, does not include MAGE 4 or MEG-E8 proteins, or consists of MAGE D protein or its immunologically active fragment. Also included are a polynucleotide encoding the immunogen or its complement, a vector comprising the polynucleotide, a mammalian cell comprising the vector and expressing the polynucleotide, a vaccine composition comprising the immunogen and an antibody specific for the immunogen. The immunogen is useful for inducing a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is useful for inducing a CTL response when administered to a subject. A

CC mammalian cell that can express the immunogen, is useful for inducing a  
 CC CTL response in vitro that is specific for a tumour cell expressing HLA-1  
 CC or HLA-2. The immunogen or cell is useful for inducing CTL for treating a  
 CC subject with cancer (carcinoma, preferably colorectal carcinoma, ovarian  
 CC carcinoma, lung carcinoma and prostate carcinoma). The immunogen is also  
 CC useful for screening and diagnostic agents, for gene screening in  
 CC patients afflicted with cancer, for screening a sample for the presence  
 CC of CTLs that specifically recognise the corresponding epitopes, as a  
 CC diagnostic tool to evaluate the efficacy of the immunotherapeutic  
 CC treatments, to prepare class I MHC (major histocompatibility class)  
 CC tetramers which are utilised in conjunction with flow cytometry to  
 CC quantitate the frequency of peptide-specific CTL that are present in a  
 CC samples of lymphocytes from an individual, and for stimulating the  
 CC production of antibodies for use in passive immunotherapy, for use as  
 CC diagnostic reagents, and for use as reagents in other processes such as  
 CC affinity chromatography. The present sequence is an immunogenic epitope  
 CC of the invention derived from human signal transducer and activator  
 CC protein

XX Sequence 9 AA;

Query Match 100.0%; Score 46; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLQELNYNL 9  
 |||||  
 Db 1 KLQELNYNL 9

RESULT 2

ABG32310  
 ID ABG32310 standard; peptide; 9 AA.

XX

AC ABG32310;

XX 05-NOV-2002 (first entry)

XX HLA-A2 associated immunogenic peptide from human KIAA0158 protein.

XX Human; immunogen; epitope; HLA-A1; human leukocyte antigen; CTL;  
 KW cytotoxic Y lymphocyte; cytostatic; cancer; colorectal carcinoma;  
 KW ovarian carcinoma; lung carcinoma; prostate carcinoma; vaccine; tumour;  
 KW HLA-2; passive immunotherapy; KIAA0158.

XX Homo sapiens.

OS WO200246416-A2.

PN 13-JUN-2002.

XX 04-DEC-2001; 2001WO-US047290.

PF 04-DEC-2000; 2000US-0251022P.

XX 20-DEC-2000; 2000US-0256824P.

XX (ARGO-) ARGONEX INC.

PA Ramakrishna V, Ross M, Philip R;

XX WPI; 2002-619021/66.

XX New immunogen useful as a vaccine for inducing cytotoxic T-lymphocyte,  
 PT and for diagnosing, preventing or treating cancer e.g. ovarian carcinoma.

XX Claim 1; Page 50; 60pp; English.

XX The invention relates to an immunogen comprising an isolated polypeptide  
 CC whose amino acid sequence comprises an epitopic peptide, does not include  
 CC MAG4 4 or MFG-E8 proteins, or consists of Mage D protein or its  
 CC immunologically active fragment. Also included are a polynucleotide  
 CC encoding the immunogen or its complement, a vector comprising the  
 CC polynucleotide, a mammalian cell comprising the vector and expressing the

CC polynucleotide, a vaccine composition comprising the immunogen and an  
 CC antibody specific for the immunogen. The immunogen is useful for inducing  
 CC a cytotoxic T lymphocyte (CTL) in vitro that is specific for a tumour  
 CC cell expressing human leukocyte antigen (HLA)-A1 or A2. The immunogen is  
 CC useful for inducing a CTL response when administered to a subject. A  
 CC mammalian cell that can express the immunogen, is useful for inducing a  
 CC CTL response in vitro that is specific for a tumour cell expressing HLA-1  
 CC or HLA-2. The immunogen or cell is useful for inducing CTL for treating a  
 CC subject with cancer (carcinoma, preferably colorectal carcinoma, ovarian  
 CC carcinoma, lung carcinoma and prostate carcinoma). The immunogen is also  
 CC useful for screening and diagnostic agents, for gene screening in  
 CC patients afflicted with cancer, for screening a sample for the presence  
 CC of CTLs that specifically recognise the corresponding epitopes, as a  
 CC diagnostic tool to evaluate the efficacy of the immunotherapeutic  
 CC treatments, to prepare class I MHC (major histocompatibility class)  
 CC tetramers which are utilised in conjunction with flow cytometry to  
 CC quantitate the frequency of peptide-specific CTL that are present in a  
 CC samples of lymphocytes from an individual, and for stimulating the  
 CC production of antibodies for use in passive immunotherapy, for use as  
 CC diagnostic reagents, and for use as reagents in other processes such as  
 CC affinity chromatography. The present sequence is an immunogenic epitope  
 CC of the invention derived from the protein encoded by the human KIAA0158  
 CC gene

XX Sequence 9 AA;

Query Match 100.0%; Score 46; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLQELNYNL 9  
 |||||  
 Db 1 KLQELNYNL 9

RESULT 3

ABU03358  
 ID ABU03358 standard; protein; 9 AA.

XX AC ABU03358;

XX 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #138.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

OS WO200278524-A2.

PN 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

PR 21-MAY-2001; 2001US-0292544P.

PR 08-AUG-2001; 2001US-0310801P.

PR 01-OCT-2001; 2001US-0326370P.

PR 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCO INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,

PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.

PS Claim 10; SEQ ID NO 138; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 9 AA;

Query Match 100.0%; Score 46; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9  
 Db 1 KQELNYNL 9  
 |||||

RESULT 4

ID AAB19969 standard; protein; 171 AA.

XX AAB19969;

DT 28-MAR-2001 (first entry)

DE Human STAT-1 DNA binding domain.

XX STAT-1; signal transducer and activator of transcription 1; human;  
 KW crystal; drug screening; DNA binding domain.

XX Homo sapiens.

XX US6160092-A.

XX 12-DEC-2000.

XX 29-MAY-1998; 98US-00087465.

XX 29-MAY-1998; 98US-00087465.

XX (UVRQ) UNIV ROCKEFELLER.

XX Chen X, Darnell JE, Kuriyan J, Vinkemeier U, Zhao Y, Jeruzalmi D;  
 XX WPI; 2001-101568/11.  
 XX N-PSDB; AAA89235.

PT Novel crystal useful in drug screening assays, comprises portion of  
 PT signal transducer, activator of transcription and duplex DNA.

XX Claim 9; Col 93-96; 206pp; English.

XX The present sequence is that of the DNA binding domain of human signal  
 CC transducer and activator of transcription 1 (STAT-1), i.e. amino acids  
 CC 318-488 of the full-length protein (see AAB19962). The domain includes an

CC immunoglobulin-type fold. The invention provides a crystal of a core  
 CC portion of a STAT protein in dimer form with an 18-mer duplex DNA (see  
 CC AAA89233) that contains a binding site for the STAT dimer. The core  
 CC portion comprises a coiled-coil domain, the DNA binding domain, a SH2  
 CC domain and a linker domain that joins the DNA binding and SH2 domains.  
 CC The crystal is of sufficient quality to perform X-ray crystallography.  
 CC studies. Methods of preparing the crystals are included in the invention.  
 CC Knowledge of the STAT protein's 3-dimensional structure will aid in  
 CC structure-based drug design. The crystal can be used in drug screening  
 CC assays to identify agonist and antagonist compounds. Antagonists can be  
 CC used to treat inflammation, allergy, asthma and leukaemia, and agonists  
 CC to treat anaemia, neutropenia, thrombocytopenia, cancer, obesity, viral  
 CC diseases, growth retardation, and other conditions characterized by  
 CC insufficient STAT activity. Fusion proteins comprising a portion of STAT,  
 CC especially the DNA binding domain, and a fusion partner are also  
 CC disclosed

XX Sequence 171 AA;

Query Match 100.0%; Score 46; DB 4; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9  
 Db 34 KQELNYNL 42  
 |||||

RESULT 5

ID ABU04889 standard; protein; 171 AA.

XX ABU04889;

DT 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1555.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.

XX Example 2; SEQ ID NO 1555; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a

CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 171 AA;

Query Match 100.0%; Score 46; DB 6; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KLOELNYNL 9  
 Db 34 KLOELNYNL 42

RESULT 6  
 AAY72844  
 ID AAY72844 standard; protein; 268 AA.  
 XX  
 AC AAY72844;  
 XX  
 DT 31-MAY-2001 (first entry)  
 XX  
 DE Mouse Stat1 protein fragment #2 (107-374 amino acids).  
 XX  
 KW Mouse; Stat1 protein; transcription factor; c-Jun; gene transcription;  
 KW cellular transformation; dysproliferative disease; cancer; psoriasis;  
 KW therapy.  
 XX  
 OS Mus musculus.  
 XX  
 FN WO200116605-A2.  
 XX  
 PD 08-MAR-2001.  
 XX  
 PF 30-AUG-2000; 2000WO-US023822.  
 XX  
 PR 31-AUG-1999; 99US-00387418.  
 XX  
 PA (UYRQ ) UNIV ROCKEFELLER.  
 XX  
 PI Zhang X, Horvath C, Wrzeszczynska MH, Darnell JE;  
 XX  
 DR WPI; 2001-226705/23.  
 XX  
 PT Identifying an agent for use in modulating the interaction between  
 PT transcription factor c-Jun and a Stat3 protein.  
 XX  
 PS Claim 65; Page 70-71; 86pp; English.

XX The present sequence is mouse Stat1 protein fragment containing 107-374  
 CC amino acids of Stat1 protein. The invention relates to methods for  
 CC identifying interacting regions of transcription factors and methods for  
 CC identifying agents which modulates the interaction between a  
 CC transcription factor such as c-Jun and a Stat protein such as Stat-1 and  
 CC Stat-3, useful for modulating gene transcription e.g., cellular  
 CC dysproliferative diseases and also for treating cancer and psoriasis. A  
 CC Stat protein comprises the N-terminal domain, coiled-coil domain, DNA  
 CC binding domain, linker domain, SH2 domain and transactivation domain

XX  
 SQ Sequence 268 AA;  
 Query Match 100.0%; Score 46; DB 4; Length 268;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KLOELNYNL 9  
 Db 244 KLOELNYNL 252  
 RESULT 7  
 AAW62996  
 ID AAW62996 standard; protein; 582 AA.  
 XX  
 AC AAW62996;  
 XX  
 DT 24-NOV-1998 (first entry)  
 XX  
 DE Human truncated Stat1 protein Stat1tc.  
 XX  
 KW Stat1 protein; signal transducer and activator of transcription; human;  
 KW purification; inflammation; allergy; asthma; leukaemia; anaemia;  
 KW neutropenia; thrombocytopenia; cancer; obesity; growth retardation;  
 KW drug screening; therapy; Stat1tc.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 24  
 FT /note= "optionally alkylated"  
 FT Modified-site 309  
 FT /note= "optionally alkylated"  
 FT Modified-site 361  
 FT /note= "optionally alkylated"  
 FT Modified-site 570  
 FT /label= O-phosphorylated  
 XX  
 FN CA2218456-A.  
 XX  
 PD 15-APR-1998.  
 XX  
 PF 15-OCT-1997; 97CA-02218456.  
 XX  
 PR 15-OCT-1996; 96US-0028176P.  
 XX  
 PA (UYRQ ) UNIV ROCKEFELLER.  
 XX  
 PI Vinkemeier U, Darnell JE;  
 XX  
 DR WPI; 1998-399711/35.  
 DR N-PSDB; AAV42537.  
 XX  
 PT New Stat protein and N-terminal fragment used in drug screening - to  
 PT obtain products for treating e.g. inflammation, allergy, asthma, anaemia,  
 PT cancer, obesity, viral disease and growth retardation.  
 XX  
 PS Claim 1; Page 82-84; 111pp; English.  
 XX  
 CC This is the amino acid sequence of a soluble, truncated, protease-  
 CC resistant Stat1 protein, designated Stat1tc, that comprises amino acid  
 CC residues 132-713 of human Stat1-alpha protein (see AAW62994). It does not  
 CC include the N-terminal compact domain of the full-length protein but  
 CC retains the Tyr701 phosphorylation site. The invention describes methods  
 CC of producing milligram quantities of 3 forms of purified Stat1 protein:  
 CC Stat1-alpha and Stat1-beta (see AAW62995) from baculovirus-infected  
 CC insect cells, and Stat1tc from E. coli. In vitro phosphorylation of the  
 CC Stat proteins plus cysteine residue alkylation to prevent aggregation,  
 CC coupled to a chromatography protocol, has allowed the purification of  
 CC activated Stat proteins. The Stat proteins and fragments can be used to  
 CC identify antagonists and agonists of Stat function. Antagonists can be  
 CC used to treat e.g. inflammation, allergy, asthma and leukaemias, and



CC agonists can be used in the treatment of e.g. anaemias, neutropaenias,  
CC thrombocytopenia, cancer, obesity, viral diseases, growth retardation or  
CC other diseases characterised by insufficient Stat activity

XX SQ Sequence 582 AA;

Query Match 100.0%; Score 46; DB 2; Length 582;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIQELNYNL 9  
| | | | |  
Db 219 KIQELNYNL 227

RESULT 8  
ABU04748  
ID ABU04748 standard; protein; 582 AA.

XX AC ABU04748;

XX DT 29-JAN-2003 (first entry)

XX DE Human expressed protein tag (EPT) #1414.

XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX protease; protease inhibitor; transporter; cytoskeletal protein;  
XX KW receptor; transcription factor; cancer; MHC;  
XX KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX OS Homo sapiens.

XX PN WO200278524-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US009671.

XX PR 28-MAR-2001; 2001US-0279495P.

XX PR 21-MAY-2001; 2001US-0292544P.

XX PR 08-AUG-2001; 2001US-0310801P.

XX PR 01-OCT-2001; 2001US-0326370P.

XX PR 04-DEC-2001; 2001US-0336780P.

XX PR 20-FEB-2002; 2002US-0358985P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Chicx RM, Tomlinson AJ, Urban RG;

XX DR WPI; 2003-040607/03.

XX PS New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX PT cytoskeletal proteins, receptors or transcription factors), useful for  
XX PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX PT leukemia.

XX PS Example 2; SEQ ID NO 1414; 134pp; English.

XX CC The invention describes a purified polypeptide, which comprises a  
XX CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX CC transporter, cytoskeletal protein, receptor or transcription factor. The  
XX CC polypeptide is useful as an immunogenic composition for eliciting in a  
XX CC mammal an immunogenic response directed against any of the purified  
XX CC polypeptide. The purified polypeptide, or the antibody that binds to this  
XX CC polypeptide, is useful for treating cancer. The polypeptide is also  
XX CC useful for identifying compounds that binds to a naturally processed  
XX CC class I or class II MHC-binding polypeptide. The polypeptides and  
XX CC polynucleotides are particularly useful for treating or preventing  
XX CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
XX CC lymphoma or leukaemia. These are also useful for screening agents for  
XX CC treating the above mentioned diseases. This sequence represents an  
XX CC expressed protein tag (EPT) isolated from human tissue for translational

CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 582 AA;

Query Match 100.0%; Score 46; DB 6; Length 582;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIQELNYNL 9  
| | | | |  
Db 219 KIQELNYNL 227

RESULT 9  
ADH57036

XX ID ADH57036 standard; protein; 582 AA.

XX AC ADH57036;

XX DT 25-MAR-2004 (first entry)

XX DE Truncated human Stat1tc protein SeqID 3.

XX KW human; Stat1tc; Stat; transcription factor;  
XX KW signal transducers and activators of transcription; cytokine;  
XX KW growth factor; protein aggregation; drug discovery; inclusion body.

XX OS Homo sapiens.

XX PN US2003092066-A1.

XX PD 15-MAY-2003.

XX PF 17-SEP-2002; 2002US-00245120.

XX PR 02-NOV-1999; 99US-00430806.

XX PA (VINK/) VINKEMEIER U.

XX PA (DARN/) DARNELL J E.

XX PI Vinkemeier U, Darnell JE;

XX DR WPI; 2004-068840/07.

XX DR N-ESDB; ADH57038.

XX PT Novel truncated signal transducers and activators of transcription  
XX PT protein and a purified N-terminal peptide fragment of the protein useful  
XX PT for drug discovery.

XX PS Claim 1; SEQ ID NO 3; 55pp; English.

XX CC This invention relates to novel methods of purifying recombinant Stat  
XX CC proteins, as well as modified and functional fragments thereof.  
XX CC Specifically, it refers to the protein family of transcription factors  
XX CC identified as the signal transducers and activators of transcription  
XX CC (Stat) that mediate the action of signalling molecules including  
XX CC cytokines and growth factors. The present invention describes various  
XX CC methods for producing milligram quantities of different forms of purified  
XX CC Stat proteins from recombinant DNA constructs, which are produced in  
XX CC insect cells infected with recombinant baculovirus. As such, it provides  
XX CC for the isolation of active tyrosine phosphorylated Stat proteins, as  
XX CC well as unphosphorylated Stat, truncated Stat proteins and those that  
XX CC exhibit modifications of specific cysteine residues that prevents protein  
XX CC aggregation. These recombinant Stat proteins are useful in the drug  
XX CC discovery process and this method that expresses the truncated protein in  
XX CC a soluble form overcomes earlier failures whereby recombinant Stat  
XX CC proteins were almost exclusively formed as insoluble inclusion bodies.  
XX CC This polypeptide sequence is the truncated human Stat1tc protein of the  
XX CC invention.

XX SQ Sequence 582 AA;

Query Match 100.0%; Score 46; DB 8; Length 582;  
Best Local Similarity 100.0%; Pred. No. 6.6; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 1 KQELNYNL 9  
| | | | | | | | |  
Db 219 KQELNYNL 227

RESULT 10  
ID ABR59713 standard; protein; 680 AA.  
XX ABR59713;  
XX  
XX 25-JUL-2003 (first entry)  
XX Human signal transducer and activator of transcription 1 isoform alpha.  
XX Human; T lymphocyte activation; T-cell; A-raf-1; TCPTP/PTPN2; asthma;  
XX immunosuppressive; antiasthmatic; antiallergic; antiinflammatory;  
XX lymphocyte activation; lymphocyte migration; cytokine production;  
XX cell surface marker expression; antibody production; apoptosis; allergy;  
XX antibody proliferation; antibody differentiation; hypersensitivity;  
XX graft versus host disease; inflammation; transducer;  
XX transcriptional activator.  
XX  
XX Homo sapiens.  
XX WO2003029277-A2.  
XX  
XX 10-APR-2003.  
XX  
XX 02-OCT-2002; 2002WO-US031618.  
XX  
XX 03-OCT-2001; 2001US-0327212P.  
XX  
XX (RIGE-) RIGEL PHARM INC.  
XX  
XX Chu P, Li C, Liao XC, Masuda E, Pardo J, Zhao H;  
XX WPI: 2003-363276/34.  
XX N-PSDB; ACC81108.  
XX  
XX Identifying a compound that modulates T lymphocyte activation, useful for  
XX monitoring changes in cell surface marker expression, comprises  
XX contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with  
XX a compound.  
XX  
XX Disclosure; Page 88; 126pp; English.

CC The invention relates to a novel method for identifying a compound that  
CC modulates T lymphocyte activation. The method comprises contacting a T  
CC cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with a compound,  
CC where the A-raf-1 or TCPTP/PTPN2 polypeptide is encoded by a nucleic  
CC acid that hybridises to a nucleic acid encoding a polypeptide having a  
CC sequence selected from two 606-amino acid sequence and a 415-amino acid  
CC sequence given in the specification. The method of the invention has  
CC immunosuppressive, antiasthmatic, antiallergic, and antiinflammatory  
CC activity. The method is useful for identifying compounds that modulate  
CC lymphocyte activation and migration, and for monitoring changes in cell  
CC surface marker expression, cytokine production, antibody production,  
CC proliferation and differentiation, and apoptosis, using either cell lines  
CC or primary cells. The A-raf-1 or TCPTP/PTPN2 proteins may be used as  
CC drug targets for compounds that suppress or activate lymphocyte  
CC activation and migration, e.g. for the treatment of diseases in which  
CC modulation of the immune response is desired such as delayed type  
CC hypersensitivity reactions, asthma, allergies, graft versus host disease,  
CC and acute and chronic inflammation. Modulators of lymphocyte activation  
CC are useful for treating disorders related T and B cell activation and  
CC migration. The present sequence is used in the exemplification of the  
CC invention

XX SQ Sequence 680 AA;  
Query Match 100.0%; Score 46; DB 6; Length 680;  
Best Local Similarity 100.0%; Pred. No. 7.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KQELNYNL 9  
| | | | | | | | |  
Db 350 KQELNYNL 358

RESULT 11  
ID AAR41335 standard; protein; 701 AA.  
XX AAR41335;  
XX  
XX 25-MAR-2003 (revised)  
XX 22-APR-1994 (first entry)  
XX 84 kD ISGF-3alpha.  
XX  
XX 113 kD; 91 kD; 84 kD; ISGF-3alpha; interferon-related; receptor;  
XX recognition factor; gene family; translation protein; tyrosine;  
XX DNA binding protein; interferon-gamma; hairy cell leukaemia;  
XX interferon therapy; chronic viral hepatitis; phosphorylation;  
XX adjuvant therapy; tyrosine kinase.  
XX  
XX Homo sapiens.  
XX  
XX WO9319179-A1.  
XX  
XX 30-SEP-1993.  
XX  
XX 19-MAR-1993; 93WO-US002569.  
XX  
XX 19-MAR-1992; 92US-00854296.  
XX 23-NOV-1992; 92US-00980498.  
XX  
XX (UYRQ ) UNIV ROCKEFELLER.  
XX  
XX Darnell JB, Schindler CW, Fu X, Shuai K;  
XX WPI: 1993-320745/40.  
XX N-PSDB; AAQ49166.  
XX  
XX Interferon receptor recognition factors - useful e.g. to treat viral  
XX hepatitis, hairy cell leukaemia and to potentiate interferon effects.  
XX  
XX Claim 17; Fig 3; 131pp; English.

CC The sequences given in AAR41333-35 represent the 113 kD, 91 kD and 84 kD  
CC ISGF-3alpha proteins respectively. ISGF-3alpha is an interferon- related  
CC receptor recognition factor which comprises several substituents. The 113  
CC kD and the 91 and 84 kD proteins are derived from two different but  
CC related genes. It is clear that a gene family exists and further members  
CC are likely to be found. The 91 kD protein has the capability of acting as  
CC a translation protein and as a DNA binding protein in response to  
CC interferon-gamma stimulation. These proteins participate in rapid  
CC phosphorylation and dephosphory- lation during the course of, and as part  
CC of their activity. This phosphorylation takes place in an interferon-  
CC dependant manner on specified tyrosine residues. These proteins may be  
CC used in conjunction with interferon therapy eg. to treat chronic viral  
CC hepatitis, hairy cell leukaemia and for use with interferon in adjuvant  
CC therapy. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 701 AA;  
Query Match 100.0%; Score 46; DB 2; Length 701;  
Best Local Similarity 100.0%; Pred. No. 8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 KLOELNYNL 9
DB      339 KLOELNYNL 347

RESULT 12
ABU04741
ID ABU04741 standard; protein; 701 AA.
XX
AC ABU04741;
XX
XX
XX 29-JAN-2003 (first entry)
XX
XX Human expressed protein tag (EPT) #1407.
XX
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
XX Homo sapiens.
OS
XX
XX WO200278524-A2.
XX
XX 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
PR
XX 21-MAY-2001; 2001US-0292544P.
PR
XX 08-AUG-2001; 2001US-0310801P.
PR
XX 01-OCT-2001; 2001US-0326370P.
PR
XX 04-DEC-2001; 2001US-0336780P.
PR
XX 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
XX Example 2; SEQ ID NO 1407; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 701 AA;

Query Match      100.0%; Score 46; DB 6; Length 701;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KLOELNYNL 9
DB      339 KLOELNYNL 347

RESULT 13
AAR72079
ID AAR72079 standard; protein; 712 AA.
XX
XX AAR72079;
XX
XX 25-MAR-2003 (revised)
DT 27-SEP-1995 (first entry)
XX
XX Human Stat84.
XX
XX Signal transducer and activator of transcription; ISGF-3; STAT; Stat84;
KW receptor recognition factor; transcription factor; cellular debilitation;
KW derangement; dysfunction; interferon-gamma.
XX
XX Homo sapiens.
OS
XX
XX WO9508629-A1.
XX
XX 30-MAR-1995.
XX
XX 26-SEP-1994; 94WO-US010849.
XX
XX 24-SEP-1993; 93US-00126588.
PR
XX 24-SEP-1993; 93US-00126595.
PR
XX 11-MAR-1994; 94US-00212184.
PR
XX 11-MAR-1994; 94US-00212185.
XX
XX (UYRQ) UNIV ROCKEFELLER.
XX
XX Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;
XX
XX WPI; 1995-139598/18.
DR
XX N-PSDB; AAQ89337.
XX
XX Receptor recognition factor implicated in transcriptional stimulation of
PT genes - useful in drug screening assays and/or for treating cellular
PT debilitations, derangements and/or dysfunctions, etc.
XX
XX Disclosure; Page 91-94; 160pp; English.
XX
XX The sequences of cDNA encoding receptor recognition factors having mol. wt.
CC of 113 kDa (Stat113), 91 kDa (Stat92) and 84 kDa (Stat84) are given in
CC AAQ89335-37 and the deduced amino acid sequences of the STAT proteins in
CC AAR72077-79. These ISGF-3-derived proteins are activated by binding of
CC interferon-alpha (all 3 Stat proteins) or interferon-gamma (Stat91) to
CC cell receptors. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 712 AA;

Query Match      100.0%; Score 46; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KLOELNYNL 9
DB      350 KLOELNYNL 358

RESULT 14
AAW03170
ID AAW03170 standard; protein; 712 AA.
XX
XX AAW03170;
XX
XX 24-OCT-1996 (first entry)
DT
XX Human STAT1-beta.
DE
```

```

XX STAT; STAT1-beta; signal transducer and activator of transcription;
KW DNA binding protein; ligand; receptor; oncogenesis; inflammation;
KW autoimmune disease; antagonist; therapy; STAT84.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 399..508
FT /label= DNA binding_domain
FT /note= "Claim 3, page 110"
XX
XX W09620954-A2.
XX
XX 11-JUL-1996.
XX
XX 28-DEC-1995; 95WO-US017025.
XX
XX 06-JAN-1995; 95US-00369796.
XX
XX (UYRQ ) UNIV ROCKEFELLER.
XX
XX Darnell JE, Wen Z, Horvath CM, Zhong Z;
XX WPI; 1996-333941/33.
XX N-PSDB; AAT31277.
XX
XX New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,
FT preventing or treating cellular dysfunction, e.g. oncogenesis,
FT inflammation, parasitic disease or autoimmunity.
XX
XX Disclosure; Page 76-79; 138pp; English.
XX
XX Signal transducer and activator of transcription (STAT) protein STAT1-
CC beta (AAW03170), also known as STAT84, is a 84 kDa protein having a dual
CC purpose, i.e. signal transduction from ligand-activated receptor kinase
CC complexes followed by nuclear translocation and DNA binding to activate
CC transcription. Recombinant STAT1-beta can be obtained using an isolated cDNA
CC clone (AAT31277). STAT1-beta includes a DNA-binding domain (see also
CC AAW03165) capable of both receptor recognition and message delivery via
CC DNA binding in a receptor-ligand specific manner. STAT1-beta is a
CC truncated form of STAT1-alpha (AAW03168). STAT proteins and their DNA
CC binding domains (see also AAW03165-69, AAW03171-76) are useful for
CC screening antagonists used to inhibit STAT-mediated signal transduction
CC and activation of transcription
XX
XX Sequence 712 AA;
XX
XX Query Match 100.0%; Score 46; DB 2; Length 712;
XX Best Local Similarity 100.0%; Pred. No. 8.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 KLOELNYNL 9
XX |||||
XX 350 KLOELNYNL 358
XX
XX RESULT 15
XX AAW62995
XX ID AAW62995 standard; protein; 712 AA.
XX
XX AC AAW62995;
XX
XX DT 24-NOV-1998 (first entry)
XX
XX DE Human Stat1-beta protein.
XX
XX KW Stat1 protein; signal transducer and activator of transcription; human;
KW purification; inflammation; allergy; asthma; leukaemia; anaemia;
KW neutropenia; thrombocytopenia; cancer; obesity; growth retardation;
KW drug screening; therapy.
XX
XX OS Homo sapiens.

```

```

XX Key Location/Qualifiers
FH Domain 1..131
FT /label= Compact_domain
FT Modified-site 155
FT Modified-site 440
FT Modified-site 492
FT Modified-site 701
FT /label= O-phosphorylated
XX
XX CA2218456-A.
XX
XX 15-APR-1998.
XX
XX 15-OCT-1997; 97CA-02218456.
XX
XX 15-OCT-1996; 96US-0028176P.
XX
XX (UYRQ ) UNIV ROCKEFELLER.
XX
XX Vinkemeier U, Darnell JE;
XX WPI; 1998-399711/35.
XX
XX New Stat protein and N-terminal fragment used in drug screening - to
FT obtain products for treating e.g. inflammation, allergy, asthma, anaemia,
FT cancer, obesity, viral disease and growth retardation.
XX
XX Claim 20; Page 80-82; 111pp; English.
XX
XX This is human Stat1-beta protein, a major proteolytic cleavage fragment
CC (aali-712) of Stat1-alpha protein (see AAW62994). Stat1-beta is resistant
CC to further proteolysis. The invention describes methods of producing
CC milligram quantities of 3 forms of purified Stat1 protein: Stat1-alpha
CC and Stat1-beta from baculovirus-infected insect cells, and truncated
CC Stat1c (see AAW62996) from E. coli. Stat1-beta has an N-terminal compact
CC domain that enhances the DNA binding of the Stat protein. The Tyr-701
CC residue is phosphorylated in vivo. In vitro phosphorylation of the
CC protein plus cysteine residue alkylation to prevent aggregation, coupled
CC to a chromatography protocol, has allowed the purification of activated
CC Stat proteins. The Stat proteins and fragments can be used to identify
CC antagonists and agonists of Stat function. Antagonists can be used to
CC treat e.g. inflammation, allergy, asthma and leukaemias, and agonists can
CC be used in the treatment of e.g. anaemias, neutropenias,
CC thrombocytopenia, cancer, obesity, viral diseases, growth retardation or
CC other diseases characterised by insufficient Stat activity
XX
XX Sequence 712 AA;
XX
XX Query Match 100.0%; Score 46; DB 2; Length 712;
XX Best Local Similarity 100.0%; Pred. No. 8.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 KLOELNYNL 9
XX |||||
XX 350 KLOELNYNL 358
XX
XX Search completed: November 18, 2005, 01:01:28
XX Job time : 48.7742 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:11:06 ; Search time 8.90323 Seconds  
(without alignments)  
97.263 Million cell updates/sec

Title: US-10-006-177-11

Perfect score: 46

Sequence: 1 KLOELNYNL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	739	2 A46159	interferon-depende
2	38	82.6	354	2 B97120	DNA uptake protein
3	36	78.3	345	2 T32203	hypothetical prote
4	36	78.3	487	2 B35059	hypothetical prote
5	36	78.3	487	2 A97928	type 1 site-specif
6	36	78.3	674	2 T28274	ORF MSV113 probabl
7	35	76.1	69	2 B33982	hypothetical prote
8	35	76.1	129	2 S41972	14.9K protein - fo
9	35	76.1	333	2 A95039	sugar binding tran
10	35	76.1	355	2 B37909	transcription regu
11	35	76.1	478	2 T33735	hypothetical prote
12	35	76.1	1005	2 A64465	hypothetical prote
13	34	73.9	120	2 F81802	probable transcrip
14	34	73.9	143	2 C81065	transcription regu
15	34	73.9	417	2 G98200	sarcosine oxidase
16	34	73.9	417	2 A33086	sarcosine oxidase
17	34	73.9	432	2 S26432	intermediate filam
18	34	73.9	567	2 T18872	intermediate filam
19	34	73.9	618	2 A30770	probable outer mem
20	34	73.9	1008	2 A32304	hypothetical prote
21	34	73.9	2573	2 D71614	hypothetical prote
22	33	71.7	180	2 C71869	hypothetical prote
23	33	71.7	216	2 A31707	L-fucose-phospha
24	33	71.7	216	2 AF1336	L-fucose-phospha
25	33	71.7	224	2 B91111	hypothetical prote
26	33	71.7	224	2 B95956	hypothetical prote
27	33	71.7	532	2 T01418	pectinesterase hom
28	33	71.7	1165	2 A46180	adenylyl cyclase t
29	33	71.7	1166	2 A49201	adenylyl cyclase

30	33	71.7	1171	2 T17454	diaphanous-related
31	32	69.6	133	2 T30474	hypothetical prote
32	32	69.6	154	2 S36993	transposase (clone
33	32	69.6	159	2 A97836	hypothetical prote
34	32	69.6	232	2 F72383	hypothetical prote
35	32	69.6	250	2 S44768	C29E4.7 protein -
36	32	69.6	261	2 G90449	conserved hypoteth
37	32	69.6	271	2 H72261	conserved hypoteth
38	32	69.6	336	2 S04682	ribosomal protein
39	32	69.6	357	2 S58215	jaomonate induced
40	32	69.6	367	2 A99468	oxidoreductase [im
41	32	69.6	375	2 D97268	toxic anion resist
42	32	69.6	420	2 S19707	site-specific DNA-
43	32	69.6	486	2 A75045	2-isopropylmalate
44	32	69.6	507	2 AG3187	aldenhyde dehydrog
45	32	69.6	528	2 E71886	type I restriction

ALIGNMENTS

RESULT 1

A46159

interferon-dependent positive-acting transcription factor ISGF-3 91K chain - human

C;Species: Homo sapiens (man)

C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C;Accession: A46159

R;Schindler, C.; Fu, X.Y.; Improt, T.; Aebersold, R.; Darnell Jr., J.E.

Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992

A;Title: Proteins of transcription factor ISGF-3: one gene encodes the 91- and 84-kDa ISG

A;Reference number: A46159; MUID:92366557; PMID:1502203

A;Accession: A46159

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: nucleic acid; protein

A;Residues: 1-739 <SCH>

A;Cross-references: UNIPROT:P42224

A;Experimental source: HeLa cells

A;Note: sequence extracted from NCBI backbone (NCBIP:110818)

C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 100.0%; Score 46; DB 2; Length 739;  
Best Local Similarity 100.0%; Pred. No. 0.58;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KLOELNYNL 9

|||||||

Db 339 KLOELNYNL 347

RESULT 2

B97120

DNA uptake protein [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004

C;Accession: B97120

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: B97120

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-354 <KUR>

A;Cross-references: UNIPROT:Q97169; GB:AE001437; PIDN:AAK79749.1; PID:gl5024755; GSPDB:G

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC1784

Query Match

Best Local Similarity 82.6%; Score 38; DB 2; Length 354;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 KLOELNYNL 9
Db 95 KJNELNYNV 103

RESULT 3
T32203
hypothetical protein T02B11.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T32203
R/Goela, D.
submitted to the EMBL Data Library, September 1997
A/Description: The sequence of C. elegans cosmid T02B11.
A/Reference number: Z21135
A/Accession: T32203
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-345 <GO>
A/Cross-references: UNIPROT:O16975; EMBL:AF022979; PIDN:AAB69906.1; GSPDB:GN000023; CESP:
A/Experimental source: strain Bristol N2; clone T02B11
C/Genetics:
A/Gene: CESP:T02B11.5
A/Map position: 5
A/Introns: 68/3; 197/2; 237/3; 294/2; 343/3

Query Match 78.3%; Score 36; DB 2; Length 345;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
Db 217 KLOKINFNL 225

RESULT 4
B95059
hypothetical protein SP0509 [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C/Accession: B95059
R/Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: B95059
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-487 <KUR>
A/Cross-references: UNIPROT:Q97S79; GB:AE005672; PIDN:AAK74667.1; PID:g14971982; GSPDB:G
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SP0509
C/Superfamily: type I site-specific deoxyribonuclease chain hsdM

Query Match 78.3%; Score 36; DB 2; Length 487;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
Db 432 ELAELNYNL 440

RESULT 5
A97928
type I site-specific deoxyribonuclease (EC 3.1.21.3) chain hsdM [imported] - Streptococ
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: A97928
```

```
R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; MUID:21429245; PMID:11544234
A/Accession: A97928
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-487 <KUR>
A/Cross-references: UNIPROT:Q8DQX1; GB:AE007317; PIDN:AAK99253.1; PID:g15458017; GSPDB:G
A/Genetics:
A/Gene: hsdM
C/Superfamily: type I site-specific deoxyribonuclease chain hsdM

Query Match 78.3%; Score 36; DB 2; Length 487;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
Db 432 ELAELNYNL 440

RESULT 6
T28274
ORF MSV113 probable early transcription factor small subunit VETF-S homolog (vaccinia D
C/Species: Melanoplus sanguinipes entomopoxvirus
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C/Accession: T28274
R/Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A/Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A/Reference number: Z20484; MUID:99102612; PMID:9847359
A/Accession: T28274
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-674 <AFO>
A/Cross-references: UNIPROT:Q9VYX9; EMBL:AF063866; NID:g4049647; PIDN:AAC97658.1; PID:g4
C/Genetics:
A/Note: MSV113
C/Superfamily: vaccinia virus early transcription factor 70K chain

Query Match 78.3%; Score 36; DB 2; Length 674;
Best Local Similarity 77.8%; Pred. No. 48;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
Db 262 KLOEDNTNM 270

RESULT 7
B83982
hypothetical protein BH2661 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: B83982
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: B83982
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-69 <STO>
A/Cross-references: UNIPROT:Q9K9I6; GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BA063
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH2661

Query Match 76.1%; Score 35; DB 2; Length 69;
```

Best Local Similarity 77.8%; Pred. No. 7.1;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 KLOELNYNL 9  
Db 36 KLSKKNYNL 44  
RESULT 8  
S41972  
14.9K protein - fowlpox virus  
C;Species: fowlpox virus  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
C;Accession: S41972  
R;Skinner, M.A.; Moore, J.B.; Binns, M.M.; Boursnell, M.E.  
submitted to the EMBL Data Library, February 1994  
A;Description: Deletion of fowlpox virus homologues of vaccinia virus genes between the  
A;Reference number: S41971  
A;Accession: S41972  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-129 <SKI>  
A;Cross-references: EMBL:Z29716; NID:G453598; PIDN:CAA82803.1; PID:G453600  
C;Superfamily: fowlpox virus 14.9K protein  
Query Match 76.1%; Score 35; DB 2; Length 129;  
Best Local Similarity 75.0%; Pred. No. 14;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 KLOELNYN 8  
Db 90 KMSLNYN 97  
RESULT 9  
A95039  
sugar binding transcription regulator RegR [imported] - Streptococcus pneumoniae (strain  
C;Species: Streptococcus pneumoniae  
C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: A95039  
R;Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A;Reference number: A95000; MUID:21357209; PMID:11463916  
A;Accession: A95039  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-333 <KUR>  
A;Cross-references: UNIPROT:Q97SK3; GB:AE005672; PIDN:AAK74506.1; PID:G14971804; GSPDB:C  
A;Experimental source: strain TIGR4  
C;Genetics:  
A;Gene: SP0330  
Query Match 76.1%; Score 35; DB 2; Length 333;  
Best Local Similarity 75.0%; Pred. No. 36;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 LOELNYNL 9  
Db 260 IKELNYNL 267  
RESULT 10  
B97909  
transcription regulator, member of GalR family regR [imported] - Streptococcus pneumoniae  
C;Species: Streptococcus pneumoniae  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C;Accession: B97909  
R;Hokins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

Y. P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A;Reference number: A97872; MUID:21429245; PMID:11544234  
A;Accession: B97909  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-355 <KUR>  
A;Cross-references: UNIPROT:Q8DR72; GB:AE007317; PIDN:AAK99102.1; PID:G15457851; GSPDB:C  
C;Genetics:  
A;Gene: regR  
Query Match 76.1%; Score 35; DB 2; Length 355;  
Best Local Similarity 75.0%; Pred. No. 39;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 LOELNYNL 9  
Db 282 IKELNYNL 289  
RESULT 11  
T33735  
hypothetical protein H34C03.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T33735  
R;Kalicki, J.; Gibson, A.  
submitted to the EMBL Data Library, October 1998  
A;Description: The sequence of C. elegans cosmid H34C03.  
A;Reference number: Z21394  
A;Accession: T33735  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-478 <KAL>  
A;Cross-references: UNIPROT:Q9TYV7; EMBL:AF100662; PIDN:AAC68976.1; GSPDB:GN00022; CESP:  
A;Experimental source: strain Bristol N2; clone H34C03  
C;Genetics:  
A;Gene: CESP:H34C03.1  
A;Map position: 4  
A;Introns: 57/1; 103/2; 185/3; 283/2; 347/3; 426/3  
C;Superfamily: Caenorhabditis elegans hypothetical protein H34C03.1  
Query Match 76.1%; Score 35; DB 2; Length 478;  
Best Local Similarity 77.8%; Pred. No. 53;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 KLOELNYNL 9  
Db 226 RLQELMYNL 234  
RESULT 12  
A64465  
hypothetical protein MJ1322 - Methanococcus jannaschii  
C;Species: Methanococcus jannaschii  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C;Accession: A64465  
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A;Reference number: A64300; MUID:96337999; PMID:8688087  
A;Accession: A64465  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1005 <BUL>  
A;Cross-references: UNIPROT:Q58718; GB:U67572; GB:L77117; NID:G1591958; PIDN:AAB99331.1;  
C;Genetics:  
A;Map position: REV1273394-1270377

C;Superfamily: hypothetical protein M1322

Query Match 76.1%; Score 35; DB 2; Length 1005;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KQLNLYN 8  
|:|||||  
Db 778 KLDEINYN 785

# RESULT 13

F81802 Probable transcription regulator NMA1774 [imported] - Neisseria meningitidis (strain Z24)

C;Species: Neisseria meningitidis  
C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C;Accession: F81802  
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A;Reference number: A81775; MUID:20222556; PMID:10761919  
A;Accession: F81802  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-120 <PAR>  
A;Cross-references: UNIPROT:Q9JTH6; GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB8500  
A;Experimental source: serogroup A, strain Z2491  
C;Genetics:  
A;Gene: NMA1774

Query Match 73.9%; Score 34; DB 2; Length 120;  
Best Local Similarity 85.7%; Pred. No. 20;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QELNLYN 9  
|:|||||  
Db 25 QDLNLYN 31

# RESULT 14

C81065 transcription regulator, MarR family NMB1585 [imported] - Neisseria meningitidis (strain

C;Species: Neisseria meningitidis  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C;Accession: C81065  
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A;Reference number: A81000; MUID:20175755; PMID:10710307  
A;Accession: C81065  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-143 <TET>  
A;Cross-references: UNIPROT:Q9JYH5; GB:AE002509; GB:AE002098; NID:g7226832; PIDN:AAF4193  
A;Experimental source: serogroup B, strain MC58  
C;Genetics:  
A;Gene: NMB1585

Query Match 73.9%; Score 34; DB 2; Length 143;  
Best Local Similarity 85.7%; Pred. No. 24;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QELNLYN 9  
|:|||||  
Db 25 QDLNLYN 31

# RESULT 15

G98200

sarcosine oxidase beta chain PA5416 [imported] - Agrobacterium tumefaciens (strain C58,  
C;Species: Agrobacterium tumefaciens  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C;Accession: G98200  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: G98200  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-417 <KUR>  
A;Cross-references: UNIPROT:Q8U7Y9; GB:AE007870; PIDN:AAK89129.1; PID:g15158937; GSPDB:GN  
C;Genetics:  
A;Gene: AGK\_L\_1106  
A;Map position: linear chromosome  
C;Superfamily: sarcosine oxidase

Query Match 73.9%; Score 34; DB 2; Length 417;  
Best Local Similarity 85.7%; Pred. No. 72;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QELNLYN 9  
|:|||||  
Db 105 QELNLYN 111

Search completed: November 18, 2005, 01:09:44  
Job time : 9.00323 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2005, 22:04:26 ; Search time 39.9677 Seconds  
(without alignments)  
115.311 Million cell updates/sec

Title: US-10-006-177-11

Perfect score: 46

Sequence: 1 KLOELNYNL 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	712	2 Q6P6Q7	Q6P6Q7 rattus norv
2	46	100.0	712	2 Q99K94	Q99K94 mus musculus
3	46	100.0	749	2 Q8C3V4	Q8C3V4 mus musculus
4	46	100.0	749	2 Q8C497	Q8C497 mus musculus
5	46	100.0	749	2 Q9D323	Q9D323 mus musculus
6	46	100.0	749	2 Q9QXK0	Q9QXK0 rattus norv
7	46	100.0	750	1 STAL HUMAN	P42224 homo sapien
8	46	100.0	755	2 Q8CD00	Q8CD00 homo sapien
9	46	100.0	755	2 Q8C8M3	Q8C8M3 mus musculus
10	46	100.0	1165	2 Q7TP57	Q7TP57 rattus norv
11	42	91.3	757	2 Q764M5	Q764M5 mus scrofa
12	40	87.0	751	2 Q8JGNO	Q8JGNO xenopus lae
13	38	82.6	354	2 Q97169	Q97169 clostridium
14	36	78.3	115	2 Q8JMC3	Q8JMC3 mamestra co
15	36	78.3	117	2 Q8QLK0	Q8QLK0 mamestra co
16	36	78.3	117	2 Q71A8H	Q71A8H mamestra co
17	36	78.3	154	2 Q9J8C0	Q9J8C0 spodoptera
18	36	78.3	175	2 Q8B102	Q8B102 mus musculus
19	36	78.3	197	2 Q8N124	Q8N124 homo sapien
20	36	78.3	269	2 Q8CFH8	Q8CFH8 mus musculus
21	36	78.3	277	2 Q87HQ0	Q87HQ0 vibrrio para
22	36	78.3	285	2 Q8XN77	Q8XN77 clostridium
23	36	78.3	399	2 Q8TBR2	Q8TBR2 homo sapien
24	36	78.3	441	2 Q8N8L3	Q8N8L3 homo sapien
25	36	78.3	487	2 Q97S79	Q97S79 streptococc
26	36	78.3	487	2 Q8DQX1	Q8DQX1 streptococc
27	36	78.3	538	2 Q95K16	Q95K16 macaca fasc
28	36	78.3	630	2 Q6CKZ4	Q6CKZ4 kluyveronyc
29	36	78.3	634	2 Q95JX7	Q95JX7 macaca fasc
30	36	78.3	674	1 E7F1 MSEPV	Q9YVX9 melanoplus
31	36	78.3	689	2 Q9FMH0	Q9FMH0 arabidopsis

#### RESULT 1

ID	Q6P6Q7	PRELIMINARY;	PRT;	712 AA.
AC	Q6P6Q7;			
DT	05-JUL-2004	(Tremblrel. 27, Created)		
DT	05-JUL-2004	(Tremblrel. 27, Last sequence update)		
DT	05-JUL-2004	(Tremblrel. 27, Last annotation update)		
DE	Stat1 protein.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
EX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzyszinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[2]			
RC	TISSUE=Prostate;			
RA	Strausberg R.;			
RL	Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC620279; AA62079.1; -			
DR	GO; GO:0005634; C:nucleus; IEA.			
DR	GO; GO:0004871; F:signal transducer activity; IEA.			
DR	GO; GO:0003700; F:transcription factor activity; IEA.			
DR	GO; GO:0007242; P:intracellular signaling cascade; IEA.			
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.			
DR	InterPro; IPR008967; PS3_like_DNA_bind.			
DR	InterPro; IPR001217; STAT.			
DR	Pfam; PF00017; SH2; 1.			
DR	Pfam; PF01017; STAT_alpha; 1.			
DR	Pfam; PF02864; STAT_bind; 1.			
DR	Pfam; PF02865; STAT_int; 1.			

Q6ZRI3 homo sapien  
O13131 oncorhynchu  
Q90Y17 tetraodon f  
Q90Y16 tetraodon f  
Q691R0 arabidopsis  
Q691W1 homo sapien  
Q691F3 homo sapien  
Q9K916 bacillus ha  
Q691H1 borrelia ga  
Q9HKJ7 thermoplasma  
Q71HX9 lactobacill  
Q69502 bartonella  
O82095 ceratopteris  
Q94Z17 pyralisella 1

#### ALIGNMENTS

```
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 712 AA; 83110 MW; 193C7E1A1DBECBCC CRC64;

Query Match 100.0%; Score 46; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KQELNYNL 9
Db 350 KQELNYNL 358

RESULT 2
Q99K94 PRELIMINARY; PRT; 712 AA.
AC Q99K94;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Stat1 protein (Mus musculus adult male spinal cord cDNA, RIKEN full-
length enriched library, clone:A330045H06 product:signal transducer
and activator of transcription 1, full insert sequence).
GN Name=Stat1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton K., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Hellon E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Jones S.J., Marra M.A.;
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska J., Schmutz J., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RA Strausberg R.;
RA Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44 (1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690 (2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
```

```
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630 (2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771 (2000).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanganaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imoto K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC004808; AA04808.1; -
DR EMBL; AK039458; BAC30355.1; -
DR HSSP; P42224; 1BF5.
DR MGD; MGI:103063; Stat1.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 712 AA; 83106 MW; DG2081709638CDA5 CRC64;

Query Match 100.0%; Score 46; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KQELNYNL 9
Db 350 KQELNYNL 358

RESULT 3
Q8C3V4 PRELIMINARY; PRT; 749 AA.
ID Q8C3V4
```



```

RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsuura H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Kasukawa T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sagaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK082706; BAC38579.1; -.
DR HSSP; P42224; 1BF5.
DR MGD; MGI:103063; Stat1.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
SQ SEQUENCE 749 AA; 87310 MW; DJF16611FF729D79 CRC64;

Query Match 100.0%; Score 46; DB 2; Length 749;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQLELYNVL 9
Db 350 KQLELYNVL 358

RESULT 5
Q9D323 PRELIMINARY; PRT; 749 AA.
AC Q9D323;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male colon cDNA, RIKEN full-length enriched
DE library, clone:903061B09 product:signal transducer and activator of
DE transcription 1, full insert sequence.
GN Name=Stat1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;

```

```

RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA The PANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsuura H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK018544; BAB31265.1; -.
DR HSSP; P42224; 1BF5.
DR MGD; MGI:103063; Stat1.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
SQ SEQUENCE 749 AA; 87280 MW; F7D40AC9D9ED7C99 CRC64;

Query Match 100.0%; Score 46; DB 2; Length 749;

```

```

Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9
Db 350 KQELNYNL 358

RESULT 6
O9QXK0
ID O9QXK0 PRELIMINARY; PRT; 749 AA.
AC O9QXK0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Signal transducer and activator of transcription 1.
GN Name=Stat1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. DOI=10.1074/jbc.M008330200;
RX Chen G., Hohmeier H.E., Newgard C.B.;
RT "Expression of the transcription factor STAT-1 alpha in insulinoma
RT cells protects against cytotoxic effects of multiple cytokines."
RL J. Biol. Chem. 276:766-772(2001).
DR EMBL; AF205604; AAF20200.1; --
DR HSPF; P42224; 1BP5.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P33_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SMO0252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 749 AA; 87234 MW; D37C634215DED355 CRC64;

Query Match 100.0%; Score 46; DB 2; Length 749;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9
Db 350 KQELNYNL 358

RESULT 7
STAL_HUMAN
ID STAL_HUMAN STANDARD; PRT; 750 AA.
AC P42224;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and activator of transcription 1-alpha/beta
DE (Transcription factor ISGF-3 components p91/p84).
GN Name=STAT1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 514-524; 654-660 AND 667-672.
RX MEDLINE=92366557; PubMed=1502203;
RA Schindler C., Fu X.-Y., Improtta T., Aebersold R., Darnell J.E. Jr.;

```

```

RT RT "Proteins of transcription factor ISGF-3: one gene encodes the 91-and
RT 84-kDa ISGF-3 proteins that are activated by interferon alpha."
RT Proc. Natl. Acad. Sci. U.S.A. 89:7836-7839(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM BETA).
RX TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=95192056; PubMed=7885841;
RA Yan R., Qureshi S., Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "The genomic structure of the STAT genes: multiple exons in coincident
RT sites in Stat1 and Stat2."
RL Nucleic Acids Res. 23:459-463(1995).
RN [4]
RP PHOSPHORYLATION SITE TYR-701
RX MEDLINE=95386533; PubMed=7657660; DOI=10.1074/jbc.270.35.20775;
RA Quelle F.W., Thierfelder W., Witthuhn B.A., Tang B., Cohen S.,
RA Ihle J.N.;
RT "Phosphorylation and activation of the DNA binding activity of
RT purified Stat1 by the Janus protein-tyrosine kinases and the epidermal
RT growth factor receptor."
RL J. Biol. Chem. 270:20775-20780(1995).
RN [5]
RP PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.
RX MEDLINE=95354205; PubMed=7543024; DOI=10.1016/0092-8674(95)90311-9;
RA Wen Z., Zhong Z., Darnell J.E. Jr.;
RT "Maximal activation of transcription by Stat1 and Stat3 requires both
RT tyrosine and serine phosphorylation."
RL Cell 82:241-250(1995).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 136-710.
RX MEDLINE=98292180; PubMed=9630226; DOI=10.1016/S0092-8674(00)81443-9;
RA Chen X., Vinkemeier U., Zhao Y., Jeruzalmi D., Darnell J.E. Jr.,
RA Kuriyan J.;
RT "Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to
RT DNA."
RL Cell 93:827-839(1998).
RN [7]
RP VARIANT STAT1 DEFICIENCY SER-706.
RX PubMed=11452125; DOI=10.1126/science.1061154;
RA Dupuis S., Hoggan C., Fieschi C., Thomassin N., Rosenzweig S.,
RA Harris J., Holland S.M., Schreiber R.D., Casanova J.L.;
RT "Impairment of mycobacterial but not viral immunity by a germline
RT human STAT1 mutation."
RL Science 293:300-303(2001).
RN [8]
RP VARIANT STAT1 DEFICIENCY PRO-600.
RX PubMed=12590259; DOI=10.1038/ng1097;
RA Dupuis S., Jouanguy E., Al-Hajjar S., Fieschi C., Al-Mohsen I.Z.,
RA Al-Junaah S., Yang K., Chapgier A., Eidsenchenk C., Bid P.,
RA Al Ghonatum A., Tufenkeji H., Frayha H., Al-Gazlan S., Al-Rayes H.,
RA Schreiber R.D., Gresser I., Casanova J.L.;

```

RT "Impaired response to interferon-alpha/beta and lethal viral disease  
 RT in human STAT1 deficiency";  
 RL Nat. Genet. 33:388-391(2003);  
 CC -!- FUNCTION: Transcription factor that binds to the IFN-stimulated  
 CC response element (ISRE) and to the GAS element. This multiprotein  
 CC transcription factor is termed ISGF3.  
 CC -!- SUBUNIT: In response to IFN alpha/beta, three subunits (STAT1-  
 CC alpha, STAT1-beta, STAT2) of ISGF3, become phosphorylated on  
 CC tyrosine, migrate into the nucleus, and assemble into a complex  
 CC together with ISGF3 gamma (p48), a DNA-binding protein that  
 CC specifically binds to the IFN-stimulated response element. In  
 CC response to IFN gamma, STAT1 forms homodimers, that also  
 CC translocate into the nucleus to activate IFN gamma-responsive  
 CC genes. Interacts with NMI.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus  
 CC in response to phosphorylation.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Alpha; Synonyms=p91;  
 CC IsoId=p42224-1; Sequence=Displayed;  
 CC Name=Beta; Synonyms=p84;  
 CC IsoId=p42224-2; Sequence=VSP\_006282;  
 CC -!- PTM: Tyrosine phosphorylated in response to IFN-gamma, IFN-alpha,  
 CC PGIF, and EGF. Serine phosphorylation is also required for maximal  
 CC transcriptional activity (lacking in beta form).  
 CC -!- DISEASE: Defects in STAT1 are the cause of STAT1 deficiency  
 CC [MIM:600555, 209950]. Patients generally suffer from mycobacterial  
 CC or viral diseases. In the case of complete deficiency, patients  
 CC can die of viral disease.  
 CC -!- SIMILARITY: Belongs to the transcription factor STAT family.  
 CC -!- SIMILARITY: Contains 1 SH2 domain.  
 CC -!- DATABASE: NAME=Starbase; NOTE=STAT1 mutation db;  
 CC WWW="http://bioinf.uta.fi/Starbase/".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi-sib.ch/announce/>  
 CC or send an email to [license@ebi-sib.ch](mailto:license@ebi-sib.ch)).  
 CC -----  
 DR EMBL; M97935; AAB64012.1; -;  
 DR EMBL; M97936; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; BC002704; AA02704.1; -;  
 DR EMBL; U18662; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18663; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18664; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18665; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18666; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18667; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18668; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18669; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; U18670; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A46159; A46159.  
 DR PDB; 1BF5; X-ray; A=136-710.  
 DR TRANSFAC; T01492; -;  
 DR Genew; HGNC:11362; STAT1.  
 DR H-InvDB; HIX0002682; -;  
 DR MIM; 600555; -;  
 DR MIM; 209950; -;  
 DR GO; GO:0005737; C:cytoplasm; TAS.  
 DR GO; GO:0005634; C:nucleus; TAS.  
 DR GO; GO:0005062; P:hematopoietin/interferon-class (D200-domain. . .; TAS.  
 DR GO; GO:0003700; P:transcription factor activity; TAS.  
 DR GO; GO:0006919; P:caspase activation; TAS.  
 DR GO; GO:0007249; P:I-kappaB kinase/NF-kappaB cascade; TAS.  
 DR GO; GO:0000074; P:regulation of cell cycle; TAS.  
 DR GO; GO:0009613; P:response to pest/pathogen/parasite; TAS.  
 DR GO; GO:0007165; P:signal transduction; TAS.  
 DR GO; GO:0007261; P:STAT protein dimerization; TAS.  
 DR GO; GO:0006366; P:transcription from pol II promoter; TAS.

DR GO; GO:0007260; P:tyrosine phosphorylation of STAT protein; TAS.  
 DR InterPro; IPR008967; P53\_like\_DNA\_bnd.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001217; STAT.  
 DR Pfam; PF000017; SH2; 1.  
 DR Pfam; PF01017; STAT\_alpha; 1.  
 DR Pfam; PF02864; STAT\_bind; 1.  
 DR Pfam; PF02865; STAT\_int; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 KW 3D-structure; Activator; Alternative splicing;  
 KW Direct protein sequencing; Disease mutation; DNA-binding;  
 KW Nuclear protein; Phosphorylation; SH2 domain;  
 KW Transcription regulation.  
 FT DOMAIN 573 670 SH2.  
 FT MOD\_RES 701 701 Phosphotyrosine (by JAK).  
 FT MOD\_RES 727 727 Phosphoserine.  
 FT VARSPPLIC 713 750 Missing (in isoform Beta).  
 FT VARIANT 600 600 /FTId=VSP\_006282.  
 FT VARIANT 706 706 L -> P (in STAT1 deficiency; complete).  
 FT VARIANT 706 706 L -> S (in STAT1 deficiency; partial.  
 FT Susceptibility to mycobacterial but not  
 FT to viral disease. Loss of GAF and ISGF3  
 FT activation. Impairs the nuclear  
 FT accumulation of GAF but not of ISGF3 in  
 FT heterozygous cells stimulated by IFNs).  
 FT /FTId=VAR\_018266.  
 FT S->A: Decreased transcriptional  
 FT activation.  
 FT MUTAGEN 727 727  
 FT HELIX 137 179  
 FT TURN 180 181  
 FT HELIX 198 247  
 FT TURN 248 249  
 FT TURN 256 256  
 FT HELIX 257 286  
 FT TURN 290 291  
 FT HELIX 293 316  
 FT STRAND 317 321  
 Query Match 100.0%; Score 46; DB 1; Length 750;  
 Best Local Similarity 100.0%; Pred. NO. 2.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KLOELNYNL 9  
 Db 350 KLOELNYNL 358  
 RESULT 8  
 Q68D00 PRELIMINARY; PRT; 750 AA.  
 ID Q68D00  
 AC Q68D00;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein DKFZp686B04100.  
 GN Name=DKFZp686B04100;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Retina;  
 RA Ottenwälder B., Obermaier B., Deutschenbaur S., Schaipp A.,  
 RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;  
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR749636; CAHL8430.1; -;  
 DR InterPro; IPR008967; P53\_like\_DNA\_bnd.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001217; STAT.  
 DR Pfam; PF00017; SH2; 1.

```

DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR PROSITE; PS50001; SH2; 1.
KW Hypothetical protein.
SQ SEQUENCE 750 AA; 87362 MW; 07B7ASE522364BA6 CRC64;

Query Match 100.0%; Score 46; DB 2; Length 750;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9
Db 350 KQELNYNL 358

RESULT 9
Q8C8M3 PRELIMINARY; PRT; 755 AA.
AC Q8C8M3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male adrenal gland cDNA, RIKEN full-length enriched
DE library, clone:B330003h23 product:signal transducer and activator of
DE transcription 1, full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RL "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format

```

```

RT sequencing pipeline with 384 multicapillary sequencer.";
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Kato F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kotoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK046517; BAC32766.1; -.
DR HSSP; P42224; 1BF5.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 755 AA; 88100 MW; 9A06C9F7EAA7099A CRC64;

Query Match 100.0%; Score 46; DB 2; Length 755;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQELNYNL 9
Db 350 KQELNYNL 358

RESULT 10
Q7TF57 PRELIMINARY; PRT; 1165 AA.
AC Q7TF57
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ab2-131.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.F.,
RA Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325191; AAP92592.1; -.
DR HSSP; P42224; 1BF5.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 2.

```



```

DR Pfam; PF01017; STAT_alpha; 2.
DR Pfam; PF02864; STAT_bind; 2.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 2.
DR PROSITE; PS00001; SH2; 2.
SQ SEQUENCE 1165 AA; 134606 MW; 7E5C8B983E5D3DE7 CRC64;

Query Match      100.0%; Score 46; DB 2; Length 1165;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
Db 781 KLOELNYNL 789

RESULT 11
Q764M5 PRELIMINARY; PRT; 757 AA.
AC Q764M5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Signal transducer and activator of transcription 1.
GN Name=stat1; (Pig).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Uenishi H., Eguchi T., Suzuki K., Sawazaki T., Toki D., Shinkai H.,
RA Okumura N., Hamashima N., Awata T.;
RT "PDE (Pig EST Data Explorer): construction of a database for ESTs
RT derived from porcine full-length cDNA libraries.";
RL Nucleic Acids Res. 32:D484-D488(2004).
DR EMBL; AB116564; BAD06318.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
SQ SEQUENCE 757 AA; 88166 MW; B6093218F2F6A029 CRC64;

Query Match      91.3%; Score 42; DB 2; Length 757;
Best Local Similarity 88.9%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
Db 350 KLOELNYNL 358

RESULT 12
Q8JGNO PRELIMINARY; PRT; 751 AA.
AC Q8JGNO;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Stat1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.

```

```

OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21098508; PubMed=11164887; DOI=10.1016/S0145-305X(00)00050-1;
RA Turpen J.B., Carlson D.L., Huang C.;
RT "Cloning and developmental expression of Xenopus Stat1.";
RL Dev. Comp. Immunol. 25:219-229(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Turpen J.B., Carlson D.L., Huang C.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY101602; AAM51552.1; -
DR HSP; P42224; 1BF5.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
SQ SEQUENCE 751 AA; 86497 MW; 788810A08B0889EA CRC64;

Query Match      87.0%; Score 40; DB 2; Length 751;
Best Local Similarity 88.9%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
Db 350 KLOELNYNL 358

RESULT 13
Q97169 PRELIMINARY; PRT; 354 AA.
AC Q97169;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE DNA uptake protein.
GN OrderedLocNames=CAC1784;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RX DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolff Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
RL EMBL; AB007687; AAK79749.1; -
DR PIR; B97120; B97120.
DR GO; GO:0009294; P:DNA mediated transformation; IEA.
DR InterPro; IPR003488; SMF.
DR Pfam; PF02481; SMF; 1.
DR TIGRFAMs; TIGR00732; dprA; 1.
DR Complete proteome.
SQ SEQUENCE 354 AA; 40570 MW; 105FDE089BFCD8C2 CRC64;

Query Match      82.6%; Score 38; DB 2; Length 354;

```



Best Local Similarity 77.8%; Pred. No. 44;  
Matches 7; Conservative 1; Mismatches 0; Gaps 0; Indels 1; Gaps 0;

QY 1 KLQELNYNL 9  
|||  
Db 95 KLNELNYNV 103

## RESULT 14

QBUMC3  
ID Q8JMC3 PRELIMINARY; PRT; 115 AA.  
AC Q8JMC3;  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
DE Hypothetical protein.  
OS Mamestra configurata nucleopolyhedrovirus B.  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=204440;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22080411; PubMed=12083822; DOI=10.1006/viro.2002.1411;  
RA Li L., Donly C., Li Q., Willis L.G., Keddie B.A., Erlandson M.A.,  
RA Theilmann D.A.;  
RT "Identification and genomic analysis of a second species of  
RT nucleopolyhedrovirus isolated from Mamestra configurata.";  
RL Virology 297:226-244 (2002).  
DR EMBL; AY126275; AAM95016.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 115 AA; 13082 MW; 0D35176637F6A7D2 CRC64;

Query Match 78.3%; Score 36; DB 2; Length 115;  
Best Local Similarity 75.0%; Pred. No. 33;  
Matches 6; Conservative 2; Mismatches 0; Gaps 0; Indels 0;

QY 2 LOELNYNL 9  
:|||||  
Db 103 IOELNYNV 110

## RESULT 15

QBQLK0  
ID Q8QLK0 PRELIMINARY; PRT; 117 AA.  
AC Q8QLK0;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mamestra configurata nucleopolyhedrovirus (MaconPV).  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=191492;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=90/2;  
RX MEDLINE=97163493; PubMed=9010313;  
RA Li S., Erlandson M., Moody D., Gillott C.;  
RT "A physical map of the Mamestra configurata nucleopolyhedrovirus  
RT genome and sequence analysis of the polyhedrin gene.";  
RL J. Gen. Virol. 78:265-271 (1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=90/2;  
RX MEDLINE=21884635; PubMed=11886270; DOI=10.1006/viro.2001.1313;  
RA Li Q., Donly C., Li L., Willis L.G., Theilmann D.A., Erlandson M.;  
RT "Sequence and organization of the Mamestra configurata  
RT nucleopolyhedrovirus genome.";  
RL Virology 294:106-121 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=90/2;  
RA Li Q., Donly C., Li L., Willis L.G., Theilmann D.A., Erlandson M.A.;

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U59461; AAM09142.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 117 AA; 13296 MW; E9E0E3DAD6ED724A CRC64;

Query Match 78.3%; Score 36; DB 2; Length 117;  
Best Local Similarity 75.0%; Pred. No. 33;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LOELNYNL 9  
:|||||  
Db 105 IOELNYNV 112

Search completed: November 18, 2005, 01:08:27  
Job time : 40.9677 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 00:48:52 ; Search time 13.6452 Seconds  
(without alignments)  
49.237 Million cell updates/sec

Title: US-10-006-177-11  
Perfect score: 46  
Sequence: 1 KQELNYNL 9

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: Issued Patents AA.\*  
2: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pap.\*  
7: /cgn2\_6/ptodata/1/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	171	3	US-09-087-465-18
2	46	100.0	268	3	US-09-387-418A-12
3	46	100.0	582	4	US-09-430-806A-3
4	46	100.0	712	1	US-08-369-796-6
5	46	100.0	712	2	US-08-852-091-6
6	46	100.0	712	2	US-08-820-754-6
7	46	100.0	712	3	US-08-956-652-6
8	46	100.0	712	3	US-08-956-869-6
9	46	100.0	712	3	US-08-948-547-6
10	46	100.0	712	3	US-08-956-653A-6
11	46	100.0	712	4	US-08-212-185-6
12	46	100.0	712	4	US-09-430-806A-2
13	46	100.0	712	5	PCT-US95-17025-6
14	46	100.0	729	4	US-09-917-254-97
15	46	100.0	740	1	US-08-276-099A-12
16	46	100.0	740	1	US-08-781-890-12
17	46	100.0	750	1	US-08-369-796-4
18	46	100.0	750	2	US-08-852-091-4
19	46	100.0	750	2	US-08-820-754-4
20	46	100.0	750	3	US-08-956-652-4
21	46	100.0	750	3	US-08-956-869-4
22	46	100.0	750	3	US-08-948-547-4
23	46	100.0	750	3	US-09-087-465-2
24	46	100.0	750	3	US-09-364-970-1
25	46	100.0	750	3	US-09-364-970-8
26	46	100.0	750	3	US-08-956-653A-4
27	46	100.0	750	4	US-09-972-800A-2

28 46 100.0 750 4 US-08-212-185-4 Sequence 4, Appli  
29 46 100.0 750 4 US-09-430-806A-1 Sequence 1, Appli  
30 46 100.0 750 5 PCT-US95-17025-4 Sequence 4, Appli  
31 46 100.0 767 4 US-09-949-016-8350 Sequence 8350, Ap  
32 36 78.3 487 4 US-09-583-110-4658 Sequence 4658, Ap  
33 36 78.3 490 4 US-09-107-433-4640 Sequence 4640, Ap  
34 35 76.1 196 3 US-09-129-030-2 Sequence 2, Appli  
35 35 76.1 333 4 US-09-583-110-4020 Sequence 4020, Ap  
36 35 76.1 345 4 US-09-248-796A-19172 Sequence 19172, A  
37 35 76.1 361 4 US-09-107-433-5205 Sequence 5205, A  
38 34 73.9 212 4 US-09-270-767-48155 Sequence 48155, A  
39 34 73.9 218 3 US-09-134-001C-5032 Sequence 5032, Ap  
40 34 73.9 306 4 US-09-107-532A-5918 Sequence 5918, Ap  
41 34 73.9 348 4 US-09-134-000C-5396 Sequence 5396, Ap  
42 34 73.9 660 4 US-09-248-796A-19966 Sequence 19966, A  
43 33 71.7 244 4 US-09-328-352-5597 Sequence 5597, Ap  
44 33 71.7 250 4 US-09-248-796A-27542 Sequence 27542, A  
45 33 71.7 566 4 US-09-411-628-4 Sequence 4, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-087-465-18  
; Sequence 18, Application US/09087465A  
; Patent No. 6160092  
; GENERAL INFORMATION:  
; APPLICANT: Vinkemeier, Uwe  
; APPLICANT: Chen, Xiaomin  
; APPLICANT: Darnell Jr., James E  
; APPLICANT: Kuriyan, John  
; TITLE OF INVENTION: A CRYSTAL OF THE CORE PORTION OF A STAT AND METHODS OF  
; TITLE OF INVENTION: USE  
; FILE REFERENCE: 600-1-229  
; CURRENT APPLICATION NUMBER: US/09/087.465A  
; CURRENT FILING DATE: 1998-05-29  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-087-465-18

Query Match 100.0%; Score 46; DB 3; Length 171;  
Best Local Similarity 100.0%; Pred. No. 0.2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 1 KQELNYNL 9  
Db 34 KQELNYNL 42

##### RESULT 2

US-09-387-418A-12  
; Sequence 12, Application US/09387418A  
; Patent No. 6391572  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Xiaokui  
; APPLICANT: Wrzeszczynska, Melissa H  
; APPLICANT: Horvath, Curt M  
; APPLICANT: Darnell Jr., James E  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR  
; TITLE OF INVENTION: INTERACTIONS  
; FILE REFERENCE: 600-1-253  
; CURRENT APPLICATION NUMBER: US/09/387.418A  
; CURRENT FILING DATE: 1999-08-31  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 268  
; TYPE: PRT

```
; ORGANISM: Mus musculus
US-09-387-418A-12

Query Match      100.0%; Score 46; DB 3; Length 268;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      244 KLOELNYNL 252

RESULT 3
US-09-430-806A-3
; Sequence 3, Application US/09430806A
; Patent No. 6720154
; GENERAL INFORMATION:
; APPLICANT: Vinkemeier, Uwe
; APPLICANT: Darnell Jr., James E.
; TITLE OF INVENTION: PURIFIED STAT PROTEINS AND METHODS OF PURIFYING THEREOF
; FILE REFERENCE: 600-1-182 N
; CURRENT APPLICATION NUMBER: US/09/430.806A
; CURRENT FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 08/951,130
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: 60/028,176
; PRIOR FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-430-806A-3

Query Match      100.0%; Score 46; DB 4; Length 582;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      219 KLOELNYNL 227

RESULT 4
US-08-369-796-6
; Sequence 6, Application US/08369796
; Patent No. 5716622
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,091
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/369,796
; FILING DATE: 06-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-852-091-6

Query Match      100.0%; Score 46; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 358

RESULT 5
US-08-852-091-6
; Sequence 6, Application US/08852091
; Patent No. 5883228
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,091
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/369,796
; FILING DATE: 06-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-852-091-6

Query Match      100.0%; Score 46; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-369-796-6

Query Match      100.0%; Score 46; DB 1; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-369-796-6
```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
|||  
Db 350 KLOELNYNL 358

RESULT 6  
US-08-820-754-6  
; Sequence 6, Application US/08820754  
; Patent No. 5976835  
; GENERAL INFORMATION:  
; APPLICANT: Darnell Jr., James E.  
; APPLICANT: Schindler, Christian W.  
; APPLICANT: Fu, Xian-Yuan  
; APPLICANT: Wen, Zilong  
; APPLICANT: Zhong, Zhong  
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/820.754  
; FILING DATE: 19-MAR-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,185  
; FILING DATE: 11-MAR-1994  
; APPLICATION NUMBER: US 07/980,498  
; FILING DATE: 23-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/854,296  
; FILING DATE: 19-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO US93/02569  
; FILING DATE: 19-MAR-1993  
; APPLICATION NUMBER: US 08/126,588  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-820-754-6

Query Match 100.0%; Score 46; DB 2; Length 712;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
|||  
Db 350 KLOELNYNL 358

RESULT 7  
US-08-956-652-6  
; Sequence 6, Application US/08956652  
; Patent No. 6013475  
; GENERAL INFORMATION:  
; APPLICANT: Darnell Jr., James E.  
; APPLICANT: Schindler, Christian W.  
; APPLICANT: Fu, Xian-Yuan  
; APPLICANT: Wen, Zilong  
; APPLICANT: Zhong, Zhong  
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/956.652  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,185  
; FILING DATE: 11-MAR-1994  
; APPLICATION NUMBER: US 07/980,498  
; FILING DATE: 23-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/854,296  
; FILING DATE: 19-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO US93/02569  
; FILING DATE: 19-MAR-1993  
; APPLICATION NUMBER: US 08/126,588  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-956-652-6

Query Match 100.0%; Score 46; DB 3; Length 712;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
|||  
Db 350 KLOELNYNL 358

RESULT 8  
US-08-956-869-6  
; Sequence 6, Application US/08956869

; Patent No. 6030808  
; GENERAL INFORMATION:  
; APPLICANT: Darnell Jr., James E.  
; APPLICANT: Schindler, Christian W.  
; APPLICANT: Fu, Xian-Yuan  
; APPLICANT: Wen, Zilong  
; APPLICANT: Zhong, Zhong  
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/956.869  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/212,185  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/854,296  
; FILING DATE: 19-MAR-1992  
; APPLICATION NUMBER: WO US93/02569  
; FILING DATE: 19-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/126,588  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-956-869-6

Query Match 100.0%; Score 46; DB 3; Length 712;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
Db 350 KLOELNYNL 358

RESULT 9  
US-08-948-547-6  
; Sequence 6, Application US/08948547  
; Patent No. 6124118  
; GENERAL INFORMATION:  
; APPLICANT: Darnell Jr., James E.  
; APPLICANT: Schindler, Christian W.  
; APPLICANT: Fu, Xian-Yuan  
; APPLICANT: Wen, Zilong  
; APPLICANT: Zhong, Zhong

; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/948,547  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/212,185  
; FILING DATE: 11-MAR-1994  
; APPLICATION NUMBER: US 07/980,498  
; FILING DATE: 23-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/854,296  
; FILING DATE: 19-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO US93/02569  
; FILING DATE: 19-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/126,588  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-948-547-6

Query Match 100.0%; Score 46; DB 3; Length 712;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
Db 350 KLOELNYNL 358

RESULT 10  
US-08-956-653A-6  
; Sequence 6, Application US/08956653A  
; Patent No. 6338949  
; GENERAL INFORMATION:  
; APPLICANT: Darnell Jr., James E.  
; APPLICANT: Schindler, Christian W.  
; APPLICANT: Fu, Xian-Yuan  
; APPLICANT: Wen, Zilong  
; APPLICANT: Zhong, Zhong  
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,653A  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
FILING DATE: 11-MAR-1994  
APPLICATION NUMBER: US 08/212,185  
FILING DATE: 23-NOV-1992  
APPLICATION NUMBER: US 07/980,498  
FILING DATE: 19-MAR-1993  
APPLICATION NUMBER: US 07/854,296  
FILING DATE: 19-MAR-1992  
APPLICATION NUMBER: WO US93/02569  
FILING DATE: 19-MAR-1993  
APPLICATION DATA:  
APPLICATION NUMBER: US 08/126,588  
FILING DATE: 24-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-195  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 712 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-956-653A-6

Query Match 100.0%; Score 46; DB 3; Length 712;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
Db 350 KLOELNYNL 358

RESULT 11  
US-08-212-185-6  
Sequence 6, Application US/08212185  
Patent No. 6605442  
GENERAL INFORMATION:  
APPLICANT: Darnell Jr., James E.  
APPLICANT: Schindler, Christian W.  
APPLICANT: Fu, Xian-Yuan  
APPLICANT: Wen, Zilong  
APPLICANT: Zhong, Zhong  
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/212,185  
FILING DATE: 11-MAR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
FILING DATE: 23-NOV-1992  
APPLICATION NUMBER: US 07/980,498  
FILING DATE: 19-MAR-1992  
APPLICATION NUMBER: US 07/854,296  
FILING DATE: 19-MAR-1992  
APPLICATION DATA:  
APPLICATION NUMBER: WO US93/02569  
FILING DATE: 19-MAR-1993  
APPLICATION DATA:  
APPLICATION NUMBER: US 08/126,588  
FILING DATE: 24-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 712 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-212-185-6

Query Match 100.0%; Score 46; DB 4; Length 712;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
Db 350 KLOELNYNL 358

RESULT 12

US-09-430-806A-2  
Sequence 2, Application US/09430806A  
Patent No. 6720154  
GENERAL INFORMATION:  
APPLICANT: Vinkemeier, Uwe  
APPLICANT: Darnell Jr., James E.  
TITLE OF INVENTION: PURIFIED STAT PROTEINS AND METHODS OF PURIFYING THEREOF  
FILE REFERENCE: 600-1-182 N  
CURRENT APPLICATION NUMBER: US/09/430,806A  
CURRENT FILING DATE: 1999-11-02  
PRIOR APPLICATION NUMBER: 08/951,130  
PRIOR FILING DATE: 1997-10-15  
PRIOR APPLICATION NUMBER: 60/028,176  
PRIOR FILING DATE: 1996-10-15  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 712  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-430-806A-2

Query Match 100.0%; Score 46; DB 4; Length 712;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLOELNYNL 9  
Db 350 KLOELNYNL 358

RESULT 13  
PCT-US95-17025-6  
; Sequence 6, Application PC/TUS9517025  
; GENERAL INFORMATION:  
; APPLICANT: James E. Darnell, Jr.  
; APPLICANT: Zilong Wen  
; APPLICANT: Curt M. Horvath  
; APPLICANT: Zhong Zhong  
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL  
; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/17025  
; FILING DATE: 28-DEC-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/369,796  
; FILING DATE: 06-JAN-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-17025-6

Query Match 100.0%; Score 46; DB 5; Length 712;  
Best Local Similarity 100.0%; Pred. No. 0.89;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLOELNYNL 9  
Db 350 KLOELNYNL 358

RESULT 14  
US-09-917-254-97  
; Sequence 97, Application US/09917254  
; Patent No. 6703204  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George  
; APPLICANT: Baak, Jan  
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer  
; FILE REFERENCE: B0801/7224 (JRV)  
; CURRENT APPLICATION NUMBER: US/09/917,254  
; CURRENT FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: US 60/222,093

QY 1 KLOELNYNL 9  
Db 350 KLOELNYNL 358

Query Match 100.0%; Score 46; DB 4; Length 729;  
Best Local Similarity 100.0%; Pred. No. 0.91;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLOELNYNL 9  
Db 367 KLOELNYNL 375

RESULT 15  
US-08-276-099A-12  
; Sequence 12, Application US/08276099A  
; Patent No. 5591825  
; GENERAL INFORMATION:  
; APPLICANT: McKnight, Steven L  
; APPLICANT: Hou, Jinzhao  
; TITLE OF INVENTION: INTERLEUKIN-4 SIGNAL TRANSDUCERS AND  
; TITLE OF INVENTION: BINDING ASSAYS  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/276,099A  
; FILING DATE: 15-JUL-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard Aron  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-59451-1/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 740 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-276-099A-12

Query Match 100.0%; Score 46; DB 1; Length 740;  
Best Local Similarity 100.0%; Pred. No. 0.93;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLOELNYNL 9  
Db 339 KLOELNYNL 347

Search completed: November 18, 2005, 09:36:04  
Job time : 13.6452 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2005, 01:08:37 ; Search time 44.3226 Seconds  
(without alignments)  
84.961 Million cell updates/sec

Title: US-10-006-177-11

Perfect score: 46

Sequence: 1 KLOELNYNL 9

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10F\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	9	14	US-10-006-177-10
2	46	100.0	9	14	US-10-006-177-11
3	46	100.0	9	16	US-10-473-127-138
4	46	100.0	171	16	US-10-473-127-1555
5	46	100.0	268	13	US-10-090-185-12
6	46	100.0	582	14	US-10-245-120-3
7	46	100.0	582	16	US-10-473-127-1414
8	46	100.0	701	16	US-10-473-127-1407
9	46	100.0	712	11	US-09-876-773-6
10	46	100.0	712	11	US-10-245-120-2
11	46	100.0	712	16	US-10-473-127-1401

12	46	100.0	712	16	US-10-473-127-1409
13	46	100.0	712	16	US-10-473-127-1411
14	46	100.0	712	16	US-10-473-127-1413
15	46	100.0	712	17	US-10-639-617-6
16	46	100.0	712	17	US-10-936-390-5
17	46	100.0	739	16	US-10-473-127-1406
18	46	100.0	749	9	US-09-833-205-4
19	46	100.0	750	9	US-09-833-205-2
20	46	100.0	750	11	US-09-876-773-4
21	46	100.0	750	14	US-10-245-120-1
22	46	100.0	750	14	US-10-308-279-44
23	46	100.0	750	16	US-10-755-889-352
24	46	100.0	750	16	US-10-755-889-823
25	46	100.0	750	16	US-10-473-127-1402
26	46	100.0	750	16	US-10-473-127-1403
27	46	100.0	750	16	US-10-473-127-1408
28	46	100.0	750	16	US-10-473-127-1410
29	46	100.0	750	16	US-10-473-127-1412
30	46	100.0	750	16	US-10-473-127-1415
31	46	100.0	750	16	US-10-473-127-1416
32	46	100.0	750	17	US-10-492-043-19
33	46	100.0	750	17	US-10-639-617-4
34	46	100.0	750	18	US-10-491-545A-55
35	46	100.0	750	18	US-10-631-467-658
36	46	100.0	786	9	US-09-925-297-550
37	46	100.0	786	16	US-10-473-127-1405
38	38	82.6	459	18	US-10-450-763-32710
39	36	78.3	197	15	US-10-108-260A-2644
40	36	78.3	441	15	US-10-108-260A-3815
41	36	78.3	487	17	US-10-472-928-880
42	36	78.3	490	18	US-10-617-320-4640
43	35	76.1	240	17	US-10-732-923-16841
44	35	76.1	300	15	US-10-424-599-226964
45	35	76.1	355	17	US-10-472-928-490

## ALIGNMENTS

RESULT 1  
US-10-006-177-10  
; Sequence 10, Application US/10006177  
; Publication No. US20030165513A1  
; GENERAL INFORMATION:  
; APPLICANT: Ramakrishna, Venky  
; APPLICANT: Ross, Mark  
; APPLICANT: Philip, Ramila  
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treat  
; TITLE OF INVENTION: Diagnosis of Cancer  
; FILE REFERENCE: 26747-35  
; CURRENT APPLICATION NUMBER: US/10/006,177  
; CURRENT FILING DATE: 2001-12-04  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: US/60/251,022  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: US/60/256,824  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Epitopic Peptide  
US-10-006-177-10

Query Match 100.0%; Score 46; DB 14; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9

Db 1 KLOELNYNL 9

```
RESULT 2
US-10-006-177-11
; Sequence 11, Application US/10006177
; Publication No. US20030165513A1
; GENERAL INFORMATION:
; APPLICANT: Ramakrishna, Venky
; APPLICANT: Ross, Mark
; APPLICANT: Philip, Ramila
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte-Inducing Immunogens for Prevention, Treatm
; FILE REFERENCE: 26747-35
; CURRENT APPLICATION NUMBER: US/10/006,177
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US/60/251,022
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US/60/256,824
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Epitopic Peptide
US-10-006-177-11

Query Match          100.0%; Score 46; DB 14; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
| | | | |
Db 1 KLOELNYNL 9

RESULT 3
US-10-473-127-138
; Sequence 138, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026WO1
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-138

Query Match          100.0%; Score 46; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
| | | | |
Db 1 KLOELNYNL 9

RESULT 4
US-10-473-127-1555
; Sequence 1555, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026WO1
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1555
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1555

Query Match          100.0%; Score 46; DB 16; Length 171;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
| | | | |
Db 34 KLOELNYNL 42

RESULT 5
US-10-090-185-12
; Sequence 12, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-12

Query Match          100.0%; Score 46; DB 13; Length 268;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9
| | | | |
```

```
Db      244 KLQELNYNL 252

RESULT 6
US-10-245-120-3
; Sequence 3, Application US/10245120
; Publication No. US20030092066A1
; GENERAL INFORMATION:
; APPLICANT: Vinkemeier, Uwe
; TITLE OF INVENTION: PURIFIED STAT PROTEINS AND METHODS OF PURIFYING THEREOF
; FILE REFERENCE: 600-1-182 N
; CURRENT APPLICATION NUMBER: US/10/245,120
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US/09/430,806
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 08/951,130
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: 60/028,176
; PRIOR FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-245-120-3

Query Match      100.0%; Score 46; DB 14; Length 582;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLQELNYNL 9
        |||||
Db      219 KLQELNYNL 227

RESULT 7
US-10-473-127-1414
; Sequence 1414, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1414
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-473-127-1414

Query Match      100.0%; Score 46; DB 16; Length 582;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLQELNYNL 9
        |||||
Db      219 KLQELNYNL 227

RESULT 8
US-10-473-127-1407
; Sequence 1407, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1407
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-473-127-1407

Query Match      100.0%; Score 46; DB 16; Length 701;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLQELNYNL 9
        |||||
Db      339 KLQELNYNL 347

RESULT 9
US-09-876-773-6
; Sequence 6, Application US/09876773
; Publication No. US20040058318A1
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; Schindler, Christian W.
; Fu, Xian-yuan
; Wen, Zilong
; Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/876,773
; FILING DATE: 07-Jun-2001
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,185
```

```
; FILING DATE: 11-MAR-1994
; APPLICATION NUMBER: US 07/980,498
; FILING DATE: 23-NOV-1992
; APPLICATION NUMBER: US 07/854,296
; FILING DATE: 19-MAR-1992
; APPLICATION NUMBER: WO US93/02569
; FILING DATE: 19-MAR-1993
; APPLICATION NUMBER: US 08/126,588
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 712 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-876-773-6

Query Match      100.0%; Score 46; DB 11; Length 712;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 358

RESULT 10
US-10-245-120-2
; Sequence 2, Application US/10245120
; Publication No. US20030092066A1
; GENERAL INFORMATION:
; APPLICANT: Vinkemeier, Uwe
; TITLE OF INVENTION: PURIFIED STAT PROTEINS AND METHODS OF PURIFYING THEREOF
; FILE REFERENCE: 600-1-182 N
; CURRENT APPLICATION NUMBER: US/10/245,120
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US/09/430,806
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: 08/951,130
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: 60/028,176
; PRIOR FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-245-120-2

Query Match      100.0%; Score 46; DB 14; Length 712;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 358

RESULT 11
US-10-473-127-1401
; Sequence 1401, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026WO1
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1409
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1409

Query Match      100.0%; Score 46; DB 16; Length 712;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 358

RESULT 12
US-10-473-127-1409
; Sequence 1409, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026WO1
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1409
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1409

Query Match      100.0%; Score 46; DB 16; Length 712;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 358
```

```
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026WO1
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1401
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1401
```

```
Query Match      100.0%; Score 46; DB 16; Length 712;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 358
```

```
RESULT 12
US-10-473-127-1409
; Sequence 1409, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026WO1
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1409
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1409
```

```
Query Match      100.0%; Score 46; DB 16; Length 712;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 KLOELNYNL 9
Db      350 KLOELNYNL 358
```

RESULT 13  
US-10-473-127-1411  
; Sequence 1411, Application US/10473127  
; Publication No. US20040236091A1  
; GENERAL INFORMATION:  
; APPLICANT: Zycos Inc.  
; TITLE OF INVENTION: TRANSLATIONAL PROFILING  
; FILE REFERENCE: 08191-026W01  
; CURRENT APPLICATION NUMBER: US/10/473,127  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 60/279,495  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: 60/292,544  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/310,801  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/326,370  
; PRIOR FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: 60/336,780  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/358,985  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 2041  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1411  
; LENGTH: 712  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-473-127-1411

Query Match 100.0%; Score 46; DB 16; Length 712;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
Db 350 KLOELNYNL 358  
|||||

RESULT 14  
US-10-473-127-1413  
; Sequence 1413, Application US/10473127  
; Publication No. US20040236091A1  
; GENERAL INFORMATION:  
; APPLICANT: Zycos Inc.  
; TITLE OF INVENTION: TRANSLATIONAL PROFILING  
; FILE REFERENCE: 08191-026W01  
; CURRENT APPLICATION NUMBER: US/10/473,127  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 60/279,495  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: 60/292,544  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/310,801  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/326,370  
; PRIOR FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: 60/336,780  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/358,985  
; PRIOR FILING DATE: 2002-02-20  
; NUMBER OF SEQ ID NOS: 2041  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1413  
; LENGTH: 712  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-473-127-1413

Query Match 100.0%; Score 46; DB 16; Length 712;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
Db 350 KLOELNYNL 358  
|||||

RESULT 15  
US-10-639-617-6  
; Sequence 6, Application US/10639617  
; Publication No. US20050079543A1  
; GENERAL INFORMATION:  
; APPLICANT: Darnell Jr., James E.  
; Schindler, Christian W.  
; Fu, Xian-Yuan  
; Wen, Zilong  
; Zhong, Zhong  
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN  
; SEQUENCES AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/639,617  
; FILING DATE: 12-Aug-2003  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/212,185  
; FILING DATE: 11-MAR-1994  
; APPLICATION NUMBER: US 07/980,498  
; FILING DATE: 23-NOV-1992  
; APPLICATION NUMBER: US 07/854,296  
; FILING DATE: 19-MAR-1992  
; APPLICATION NUMBER: WO US93/02569  
; FILING DATE: 19-MAR-1993  
; APPLICATION NUMBER: US 08/126,588  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq. David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 712 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-639-617-6

Query Match 100.0%; Score 46; DB 17; Length 712;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLOELNYNL 9  
Db 350 KLOELNYNL 358  
|||||

Search completed: November 18, 2005, 10:44:46

Job time : 45.3226 secs

-----